LANGUAGE, TRUTH AND ONTOLOGY

PHILOSOPHICAL STUDIES SERIES

Founded by Wilfrid S. Sellars and Keith Lehrer

Editor:

KEITH LEHRER, *University of Arizona*

Managing Editor:

LOIS DAY, *University of Arizona*

Board of Consulting Editors:

JONATHAN BENNETT, *Syracuse University*

ALLAN GIBBARD, *University of Michigan*

ROBERT STALNAKER, *Massachusetts Institute of Technology*

ROBERT G. TURNBULL, *Ohio State University*

VOLUME 51

LANGUAGE, TRUTH AND ONTOLOGY

Edited by:

KEVIN MULLIGAN

SPRINGER-SCIENCE+BUSINESS MEDIA, B.V.

Library of Congress Cataloging-in-Publication Data

Language, truth, and ontology / edited by Kevin Mulligan.
p. cm. -- (Philosophical studies series ; v. 51)
Includes three papers presented at a colloquium held in Geneva, 3-5 June 1988 and three from a colloquium held in Zinal, June 1-3, 1990.
Includes indexes.
ISBN 978-94-010-5149-1 ISBN 978-94-011-2602-1 (eBook)
DOI 10.1007/978-94-011-2602-1
1. Ontology--Congresses. 2. Language and languages--Philosophy--Congresses. 3. Truth--Congresses. I. Mulligan, Kevin, 1951-
II. Series.
BD311.L36 1992
111--dc20 91-37262

ISBN 978-94-010-5149-1

Printed on acid-free paper

All rights reserved
© 1992 Springer Science+Business Media Dordrecht
Originally published by Kluwer Academic Publishers in 1992
Softcover reprint of the hardcover 1st edition 1992
No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without written permission from the copyright owner.

In memory of Hector-Neri Castañeda

TABLE OF CONTENTS

PREFACE

All except three of the papers in this volume were presented at the colloquium on "L'Ontologie formelle aujourd'hui", Geneva, 3-5 June 1988. The three exceptions, the papers by David Armstrong, Uwe Meixner and Wolfgang Lenzen, were presented at the colloquium on "Properties", Zinal, June 1-3, 1990. It was, incidentally, at the second of these two colloquia that the European Society for Analytic Philosophy came into being.

The fathers of analytic philosophy – Moore and Russell – were in no doubt that ontology or metaphysics as well as the topics of language, truth and logic constituted the core subject-matter of their "analytic realism",[1] for the task of metaphysics as they conceived things was the description of the world.[2] And logic and ontology are indissolubly linked in the system of the grandfather of analytic philosophy, Frege. After the Golden Age of analytic philosophy – in Cambridge and Austria – opposition to realism as well as the "linguistic turn" contributed for a long time to the eclipse of ontology.[3]

Thanks in large measure to the work of some of the senior contributors to the present volume – Roderick Chisholm, Herbert Hochberg, David Armstrong and Karel Lambert – ontology and metaphysics now enjoy once again the central position they occupied some eighty years ago in the heyday of analytic philosophy. Thus three of the younger contributors to the present volume have recently published substantial treatises in these areas: Graeme Forbes, *The Metaphysics of Modality*, 1985 (Oxford: Clarendon); Peter Simons, *Parts: A Study in Ontology*, 1987 (Oxford: Clarendon); Ingvar Johansson, *Ontological Investigations*, 1989 (London: Routledge). And a fourth, Barry Smith, is the editor (with Hans Burkhardt) of the *Handbook of Metaphysics and Ontology*, which is due to appear in 1991 (Munich: Philosophia).

The papers collected here examine the credentials, formal and philosophical, of a variety of would-be denizens of the philosopher's zoo: things, events, states, changes, boundaries, parts, sums, sets, tendencies, states of affairs, worlds and objects – complete, incomplete, variable and non-existent – as well as properties – negative and disjunctive, formal and

material, repeatable and non-repeatable, mind-dependent and mind-independent. Those papers that also study the interface between the world and its representation, sentential or diagrammatic, and in particular such semantic properties of representations as designation and truth, share a concern to make the fit between logical form and the world as realistic as possible.

For financial assistance that made possible the two colloquia in Geneva and Zinal I should like to thank: the Société Académique du Valais, the Département de l'Instruction Publique du Canton du Valais, the Académie Suisse des Sciences Humaines and my own University.

I should particularly like to thank Jean-Claude Pont for his help, Lois Day for her careful and efficient labours on the manuscripts she received and Keith Lehrer for his willingness to take on the volume.

December 1990 KEVIN MULLIGAN
Department of Philosophy
University of Geneva

NOTES

[1] Cf. Bertrand Russell, "Le réalisme analytique", in Bulletin de la Société française de philosophie, 11, 1911, 53-83.

[2] Cf. Bertrand Russell, *Logic and Knowledge*, ed. R. Marsh, Unwin: London, 215.

[3] Terminology: "Metaphysics" and "Ontology" have, it seems, come to be used as near synonyms within analytic philosophy, in spite of the connotations with which the Vienna Circle used the former. "Metaphysical commitment", though, unlike "ontological commitment", does indeed convey such connotations. "Formal ontology" is much less common in English than, say, "analytic metaphysics". The former expression, or to be exact "formale Ontologie", was a favourite expression of Husserl's. In the third of his *Logische Untersuchungen* (*Logical Investigations*, London: Routledge, 1970) the expression is used to describe the purely analytic (formal) parts of what has come to be called "analytic metaphysics" (metaphysics as done by analytic philosophers, involving the method of analysis), in other words, the formal truths of mereology, of the theory of relations of more and less, of property theory, of the theory of situations, etc. At §14 of the fourth Investigation Husserl puts forward the claim that formal logic and formal ontology are distinct when he suggests that "in part, but only in part, relations of equivalence link" the analytic laws of logic and those of formal ontology.

RODERICK M. CHISHOLM

THE BASIC ONTOLOGICAL CATEGORIES

1. INTRODUCTION

In order to set forth what I take to be the most fundamental ontological categories, I will present four dichotomies – four ways of dividing things into exclusive and exhaustive subsets. In each case one of the subsets will be characterized positively and the other negatively. I will also attempt, so far as possible, to characterize the subsets in positive terms.

The dichotomies are these:

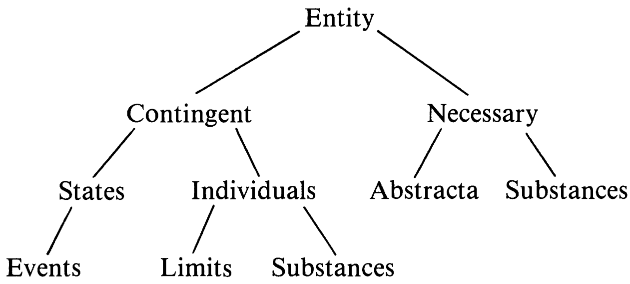
(1) Things which are *contingent* and things which are non-contingent or *necessary*;

(2) contingent things which are states and those which are non-states or *contingent individuals*;

(3) contingent individuals which are *limits* and those which are non-limits or *contingent substances*;

(4) noncontingent things which are *abstracta* and those which are non-abstracta or *noncontingent substances*.

A Table of Categories



According to this way of looking at the world, there are contingent substances along with their states and their boundaries or limits; and there are necessary things, each of which is either an abstractum or a substance. I believe we have no good reason to affirm the existence of any *other* type of thing.

2. THE BASIC CONCEPTS

Among the principal desiderata of a theory of categories are (1) economy with respect to the types of entity that are countenanced and (2) simplicity with respect to the types of concept that are used.

I will make use of these undefined concepts: (1) x *exemplifies* (has) y; (2) x is *necessarily* such that it is F; (3) x is a state of y; (4) x is a *constituent* of y; and (5) x conceives y. These concepts are all familiar. And we do not need to go beyond them in setting forth an account of the fundamental categories of ontology.

I will introduce three types of term: (1) *"being-F"*, wherein the letter "F" may be replaced by any well-formed English predicate; (2) *"x-being-F"*, which will be used to designate states of the entity designated by "x"; and (3) *"that-p"*, in which the letter "p" may be replaced by any well-formed English sentence. And I will make essential use of *tense*. For I assume that there are truths that can be adequately expressed only in a language which, like our ordinary language, is tensed.

Since I am using tensed language, I will say that whatever exists exists now. And since the language *is* tensed, the "now" in "Whatever exists exists now" is redundant and the statement is logically true. But there *was*

a philosopher who drank the hemlock. And this means that there is something – for example, the property of being blue – which *was* such as to have the property of being such that there *is* a philosopher who is drinking the hemlock.

I now turn to the four dichotomies.

(1) Contingent and Necessary Things

How are we to distinguish between those things that are contingent and those things that are not?

We have the locution, "x is necessarily such that it is F." But unfortunately we cannot use this locution to make the distinction between necessary and contingent things. For "is necessarily such that it is F," at least as I interpret it, is equivalent to "is necessarily such that it exists if and only if it is F." And if the locution is taken in this way, then *everything* may be said to "exist necessarily" – for everything is necessarily such that it exists if and only if it exists. So "exists necessarily" does not yield the distinction between necessary and contingent things.

A contingent thing, unlike a noncontingent thing, is a thing that might not have been – a thing which is possibly such that it came into being. And so I will turn to the concept of *coming into being*. (Our definition of this concept, like all the definitions presented here, is *tensed*.)

D1 x is *coming into being* = Df x is such that there is nothing it did exemplify

There are things that you and I *did* exemplify – say, the property of being a child – and therefore we are not now coming into being.

If we interpret "x is such that it is F" correctly, we will see that, if a thing is *not* possibly such that it is coming into being or passing away, then it never was and never will be possibly such that it is coming into being or passing away.

And so we draw our distinction between contingent and necessary things:

D2 x is a contingent entity = Df x is possibly such that it is coming into being

A *necessary* thing is a thing that is not contingent.

(2) States of Things

We have taken "x is a state of y" as one of our four undefined philosophical expressions. We may now make explicit the following assumption about the nature of states:

(A1) For every x, if x exemplifies being-F, then there is the state, x-being-F

Events will be exhibited as being the contingent states of certain contingent things.[1]

It will be useful to refer to the thing designated by "x" in locutions of the form "x-being-F" as being "the *substrate*" of the relevant state and to refer to the property designated by "being-F" as "the *content*" of the state. Let us say:

D3 x has being-F as its *content* and y as its *substrate* =Df (1) y is F; and (2) x is identical with y-being-F

One may object: "Aren't there *many* states of x-being-F? There is my reading now. But there are also other instances of my reading: those instances that have taken place and those that will take place." But the answer is that it is a mistake to say that there *are* those states that have taken place and that there *are* those states that will take place. There are the states that *are* taking place; there *were* those that did take place and there *will be* the ones that will take place.

Things may be said to enter into temporal and causal relations *via* their states. For example, we may say of a person that his falling contributes causally to his being injured. The cause of the injury was a "particular fall" that the person had; and the effect of the fall was a "particular injury." It was not just the *property* of falling that contributed to the injury. It was a certain particular fall; and that particular fall contributed to that particular injury. Hence there is a fall that can be distinguished from any of the subject's other falls, and a case of his being injured that can be distinguished from any other case of his being injured.

But in saying that there is "a particular fall," we are not saying that, in addition to the property of falling, there is also a kind of "particularized property" or "universal as particular." We are saying that, if a thing has the property of falling, then there is, in addition to the thing and the property, the state of *that thing having that property*.

Bernard Bolzano said that states are "beings *of* other things" and

therefore that "a state is not an *ens* per se.[2] We may express this fact by saying that a state of a thing is ontologically dependent upon its substrate – upon the thing of which it is a state. In other words:

(A2) For every x and y, if x is a state of y, then x is necessarily such that it is a state of y

States of contingent things – even the necessary states of contingent things – are themselves contingent things.

The contingent states of a thing are necessarily such that they *are* states of that thing. A person who is sad is not necessarily such that he is sad; but his-being-sad is necessarily such that it is a state of him.

3. INDIVIDUAL THINGS AND EVENTS

An individual thing is a contingent thing that is not a state:

D4 x is an *individual thing* =Df x is a contingent thing; and x is not a state of anything

We will assume a principle of *sums*, or *conjunctiva*, allowing us to say that heaps or aggregates of individuals are themselves individuals. Making use of our third undefined concept – that expressed by "x is a constituent of y" – we may put the assumption this way:

(A3) For every x and y, if x and y are individual things having no constituents in common, then there is an individual thing z such that x and y are constituents of z

(We will characterize *parts*, below, as constituting a subspecies of *constituents*.)

We may now define *events* as contingent states of contingent things. Given the concepts of content and substrate, introduced in D3 above, we may put our definition of the concept of an event this way:

D5 x is an *event* =Df (1) x is a state; (2) the substrate of x is a contingent thing; and (3) x is not such that it exemplifies its content necessarily

4. BEGINNINGS AND PROCESSES

What we have said about temporal relations should not be taken to imply that all events "exist only for an instant"; for events may endure. But

those events that are *beginnings* do "exist only for an instant." And those that we may call *"processes"* endure and therefore "exist for more than an instant.[3]

One's entire life is a single process – one that could be called one's *history* or *biography*. And, more generally, everything is such that either it is going through its history or it is beginning its history.[4] Hence we may speak of the various *stages* and *limits* of the history of a thing – or, to use more technical terms, of the *temporal parts* and *slices* of that history. But we must not make the category-mistake of supposing that such temporal parts and slices of a thing's *history* are also parts and slices of the thing that *has* the history.

By appeal to the concept of *coming into being*, that was introduced in D1, we may now characterize *beginnings* and *processes*.

D6 x is beginning to exemplify being-F =Df The state x-being-F is coming into being

D7 x is a *beginning* =Df x is a state; there is nothing that x did exemplify; and there is nothing that x will exemplify

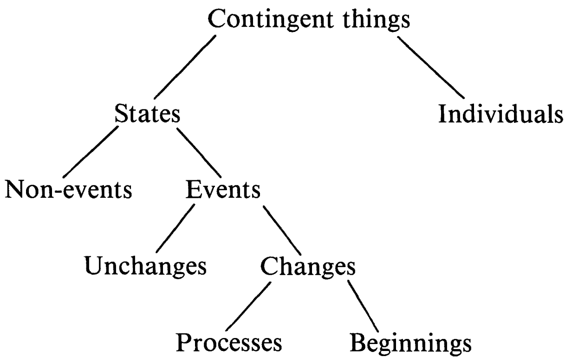
D8 x is a *process* =Df x is a state which will include a beginning

D9 x is a *change* =Df either x is a beginning or x is a process

We have characterized events as being states of contingent things. But not all beginnings and processes, and therefore not all changes, are events. For some beginnings and processes are states of non-contingent things.

Consider that situation which is an event A contributing causally to another event B. Here we have a state which has a state (event A) and not an individual as its substrate; hence it is not here counted as an "event". That situation which is one event contributing causally to another is an instance of a *state of a state* of a contingent thing and is therefore not an event but a state that we may call a "*non-event*".

The accompanying table will show how these ontological categories are related. Following C.J. Ducasse, I will use "unchanges" for those states that are not changes.[5]



There is no place in this table for "times," "facts," "tropes," or "particularized universals."

I do not assume, as many contemporary philosophers do, that there *are* such things as "times." And therefore I will not assume that there are things designated by those linguistic expressions that are *dates* – for example, "June 3, 1988." I do assume, of course, that things persist: some things *had* attributes that they no longer have; and some things *will have* attributes that they do not have. And from this it follows that some states are temporally preceded by other states and that some states are temporally followed by other states.

Our assumption according to which there are things that have come into being enables us to give a sense to such expressions as "the first time," "the second time," and "the n-th time," and to do so without supposing that there are such entities as "times."[6]

The rejection of the ontological category of *times* is quite consistent with the acceptance of *space-time*. For space-time is not an individual thing but a *state* of a very widely scattered individual thing.

(2) Limits and Contingent Substances

An adequate theory of categories should enable us to distinguish between those contingent individual things that may properly be called *substances* from those contingent things that are the *limits* or *boundaries* of substances.

To make this distinction, I use the locution, "x is a constituent of y." I assume that the relation of being-a-constituent-of is asymmetrical and

transitive: for every x and y, if x is a constituent of y, then y is not a constituent of x; and for every x, y and z, if x is a constituent of y and if y is a constituent of z, then x is a constituent of z.[7]

A *boundary* or *limit* is, in the following sense, a dependent individual:

D10 x is a *boundary* (limit) =Df (1) x is a contingent individual; and (2) every constituent of x is necessarily such that either it is a constituent or it is coming into being or passing away

Now we may say what a *contingent substance* is:

D11 x is a *contingent substance* =Df x is a contingent individual that is not a limit

The *parts* of a contingent substance are those of its constituents that are not boundaries or limits.

D12 x is a *part* of a substance y =Df x is a constituent of y; x is not a boundary of y; and y is a contingent substance

It follows, therefore, that the parts of a contingent substance are themselves contingent substances. Hence we should reject the view of Aristotle, according to which the parts of actual substances are not themselves actual substances. And we should reject the view of Leibniz, according to which actual substances cannot be parts of actual substances.

(3) Abstracta: Necessary Things that are not Individuals

It is often said that there is just one *ens necessarium* – namely God. But if, as I believe, extreme realism, or Platonism, is true, and if the distinction between contingent and necessary things is to be drawn in the way that I have suggested, then it follows that there are indefinitely many necessary things. All so-called "abstracta" are necessary things – things incapable of coming into being or passing away. These include, not only exemplified attributes, such as the attribute of being a dog, but also unexemplified attributes, such as the that of being a unicorn or that of being a round square.

Our undefined concepts include the intentional concept expressed by "x conceives y." We make use of this concept in saying what an *attribute* is:

D13 P is an *attribute* =Df (1) P is possibly such that it is conceived; and (2) either (a) P is possibly such that it is exemplified or (b)

there is a Q which is necessarily such that it is exemplified and such that whoever conceives it conceives P

The second clause enables us to say that there are such impossible attributes as being a round square. (There is a necessary attribute – being such as to be not both round and square – which is necessarily such that whoever conceives it conceives the attribute of being both round and square.)

Are there abstracta that are not attributes?

What of *classes*, or *sets*? Russell showed how the principles of set-theory may be construed as being principles about attributes. To say that x is a member of the class of F's is to say that x is F; to say that the class of F's includes the class of G's is to say that everything that is G is F; and, more generally, to say that the class of F's is so-and-so is to say that the attribute of being-an-F is exemplified by exactly the same things as is an attribute that is so-and-so.[8] Following Russell, then, we will say this:

D14 The *class* of F's is G =Df There is an attribute which is such that (a) it is G and (b) it is exemplified by all and only those things that exemplify being-F

Given that we have countenanced the being of attributes, there is no need to assume, therefore, that in *addition to* attributes there are also such things as classes or sets. Our definition, however, enables us to use the convenient terminology of classes or sets. (Since we do not suppose that there are sets in addition to attributes, we need not face such difficult questions as: "Do sets have their members necessarily?" and "Can sets change their members?")

Aren't *states of affairs* to be counted among abstract objects? To be sure, there *is* that abstract state of affairs which is all men being mortal, as well as that state of affairs which is some men not being mortal. But we may characterize these states of affairs by reference to *being such that all men are mortal* and *being such that no men are mortal*. More generally, the state of affairs that-p may be characterized by reference to *being such that p*. Our definition is this:

D15 That-p is a *state of affairs* =Df There is an attribute which is necessarily such that it is exemplified only if p

This definition guarantees that states of affairs are abstract things and not contingent events. For attributes are necessary things and therefore if an

attribute can be said to be *"necessarily* such that p," then the sentence replacing "p" cannot express a contingent event.

Since states of affairs are thus reducible to attributes, the expression "the state of affairs that-p *obtains*" is reducible to a statement about exemplification:

D16 The state of affairs that-p *obtains* =Df Something has an attribute which is necessarily such that it is exemplified if and only if p

What of *propositions*? If we use "proposition" to refer to a type of abstract object and not to a type of contingent thing (such as those "singular propositions" that are thought to contain contingent things as their constituents), then there would seem to be no ground for distinguishing propositions from states of affairs – unless we say that propositions are those states of affairs which are necessarily such that either they are always exemplified or they are never exemplified. We will say, then:

D17 That-p is a *proposition* =Df That-p is a state of affairs which is necessarily such that either it always obtains or it never obtains

The concept of the *truth* of a proposition would be explicated by reference to exemplification in the way suggested above:

D18 The proposition that-p is *true* =Df The proposition that-p obtains

To the question, "What is there *in virtue of which* any given proposition is true?", the most obvious answer would seems to be:

The *state* of some entity.[9]

"But if you assume that there are 'truth-makers' in virtue of which true propositions are true, then you should assume that there are *falsehood*-makers in virtue of which false propositions are false. How do *they* fit into your ontology." What makes a false proposition false is that truth-maker that makes it contradictory true.

If we need to speak of such things as *possible worlds*, then we may identify such entities with a certain type of state of affairs.

D19 W is a *world* =Df W is a state of affairs such that: for every state of affairs p, either W logically implies p or W logically

implies the negation of p; and there is no state of affairs q such that W implies both q and the negation of q

In other words, a world is a self-consistent, maximal state of affairs. That it is maximal is guaranteed by the first clause of the definition; and that it is self-consistent is guaranteed by the second.

If possible worlds are thus reducible to states of affairs, and if states of affairs are reducible to attributes, then possible worlds are reducible to attributes.

What of *relations*? In contemporary logic it is usual to assimilate relations to sets of a certain sort – namely, those sets that are *ordered-pairs*. For example, following Kuratowski, one may construe the ordered-pair, x-paired-with-y, as the set whose sole members are (a) the set whose sole member is x and (b) the set whose sole members are x and y. Then one is able to distinguish the set, x-paired-with-y, from the set y-paired-with-x.[10] This conception is readily carried over into the theory of attributes.

An *ordered attribute*, John-to-Mary, will be any attribute whose sole instances are: (1) an attribute whose sole instance is John and (2) an attribute whose sole instances are John and Mary. If John is, say, the tallest man in town and if John and Mary are the only people in the green Chevrolet, then an instance of the ordered attribute, John-to-Mary, would be the attribute of being exemplified only by the tallest man in town and by the only people in the green Chevrolet. *This* attribute, then, is one of those things that are exemplified by the relation, *being taller than*. And so, too, for any of the other attributes that are thus ordered from John to Mary (say, the attribute of being examplified only by the local television repairman and by the 17th and the 284th person to have registered in the last local election). *All* of these ordered attributes exemplify the relation being-taller-than. Hence we may say that John is paired to Mary by the relation *being taller than*.

Relations, then, will be characterized as follows:

D20 R is *ordered* from x to y =Df R is an attribute whose sole instances are (a) an attribute whose sole instance is x and (b) an attribute whose sole instances are x and y

D21 R is a *relation* =Df There exists an x and a y such that R is ordered from x to y

This conception of relations may be extended to relations of any number of terms. If a relation has 3 terms, then one of its instances has 2 terms and the other just one; if a relation has 4 terms, then either (a) each of its instances has 2 terms or (b) one of its instances has 3 terms and the other just one; and so on. Every relation, no matter how many terms it has, is an *attribute* – an attribute that is exemplified by other attributes.

5. NECESSARY SUBSTANCE

We will assume, then, that there are two types of non- contingent being – those that are attributes and those that are not attributes. And we will say that, if there is a necessary being that is not an attribute, then it is a *necessary substance*:

D22 x is a *necessary* substance = Df x is not contingent and x is not an attribute

A necessary substance is an eternal object. And since a necessary substance is not possibly such that there *is* anyone who attributes it to anything, we may say with Aristotle, that it is something which is "not said *of* a subject."

Such a characterization of necessary substance is essentially *negative*. We will not here consider the question whether there *is* such a being or whether, *if* there is such a being, we may characterize it in positive terms.

NOTES

[1] In holding that states and events are contingent things, I depart in a fundmental way from the view of events as abstract objects that I attempted to defend in *Person and Object* (LaSalle, IL: The Open Court Publishing Co., 1976); see Chapter IV. In arriving at the present view, I have been especially influenced by the work of Jaegwon Kim. See especially his "Events as Property Exemplifications," in M. Brand and D. Walton, eds., *Action Theory* (Dordrecht: D. Reidel, 1976), pp. 159-77.

[2] Bolzano applies the concept this way: "Everything that there is of one or the other of the following two types: either it is an entity which is *of* another thing [*an etwas Anderem*] or it exists, as one is accustomed to saying, *in itself [fur sich]*." Bernard Bolzano, *Athanasia oder Gründe für die Unsterblichkeit der Seele* (Sulzbach: J.G.V. Seidelschen Buchhandlung, 1838), p. 21.

[3] The distinction between beginnings and processes was clearly set forth by Roman Ingarden. But Ingarden, conceding that he was using terms somewhat arbitrarily, restricted "events" (*"Ereignisse"*) to what are here called "beginnings," suggesting that processes (*"Vorgänge"*) are not properly called "events." See *Der Streit um die Existenz der Welt*

(Tubingen: Max Niemeyer Verlag, 1964), Vol. I, pp. 191ff.

[4] We will not say that there are also "endings"; for "endings" may be reduced to beginnings. Instead of saying "*a* is ceasing to be," we may say: "There exists a y which is beginning to be such that *a* does not exist." I have discussed the relevant issues in *On Metaphysics* (Minneapolis: The University of Minnesota Press; 1989), see the essay "Coming Into Being and Passing Away: Can the Metaphysician Help?", pp. 49-61.

[5] See C.J. Ducasse, *Truth, Knowledge and Causation* (London: Routledge & Kegan Paul, 1968), p. 5.

[6] I have discussed these points in detail in "Events Without Times: An Essay on Ontology," forthcoming in *Nous*.

[7] Compare my "Boundaries as Independent Particulars," *Grazer Philosophische Studien*, Vol. XX (1983), pp. 87-95. This conception of boundaries comes from Brentano. Suarez had raised the question whether God might continuously destroy the parts of a cone with the result that at some moment only the point of the cone would remain in existence. The above definition presupposes that this is not possible, but the definition is readily modified to accommodate such a possibility. We have only to replace clause (2) by: "every constituent of x is necessarily such that either it is a constituent or it is coming into being or passing away."

[8] See page 249 of Russell's "Mathematical Logic as Based on the Theory of Types," *American Journal of Mathematics*, Vol. XXX (1908), pp. 222-62. This material is included in A.N. Whitehead and Bertrand Russell, *Principia Mathematica*, Vol. I (Cambridge: The University Press, 1935), pp. 71ff., and 187ff. See the discussion of Russell's definition in R. Carnap, *Meaning and Necessity* (Chicago: The University of Chicago Press, 1946), pp. 147-51.

[9] Concerning this question, see Kevin Mulligan, Peter M. Simons and Barry Smith, "Truth-makers," *Philosophy and Phenomenological Research*, Vol. 44 (1984), pp. 287-321.

[10] For example, see. W.V. Quine, *Mathematical Logic* (New York: W.W. Norton and Company, 1940), p. 198. See also Quine's *Word and Object* (Cambridge: MIT Press, 1966), pp. 257-9.

D. M. ARMSTRONG

PROPERTIES

In the present climate of metaphysics nothing is more important, I think, than the recognition of properties and relations as fundamental constituents of reality. Once properties and relations are admitted, further questions can be raised. Should we, as our languages seem to urge us, admit alongside properties and relations, things, particulars, which have the properties and between which relations hold? Or should we instead try to construct particulars out of properties and relations, or even out of properties alone, or relations alone? Again, should we take properties and relations as universals, that is, should we take it that different particulars can have the very same property, in the full strict sense of the word 'same', and that different pairs, triples .... n-tuples ....can be related by the very same relation? Or should we instead hold that properties and relations are particulars (abstract particulars, tropes, moments) so that each particular has its own properties that no other particular can have, and pairs, etc. of particulars each their own relations? A third issue: should we allow that properties and relations themselves can be propertied and stand in relations? Or should we instead with Brian Skyrms (1981) allow nothing but a first level of properties and relations?

These issues, and others, about properties and relations are of the greatest interest. And because an answer to one of the questions does not in any obvious way pre-empt the answer to any of the others, we have here a sort of metaphysician's paradise in which philosophers can wander, arguing. But before these issues can be joined there must be established the fundamental point: that there are in reality properties and relations. In this paper, I will largely confine myself to properties.

1. WHY WE SHOULD ADMIT PROPERTIES

The great deniers of properties and relations are of two sorts: those who put their faith in *predicates* and those who appeal to *sets* (classes). Some seem to take their comfort from both at once. The resort to predicates was, I suppose, given encouragement by the great Linguistic Turn, with its hope of solving philosophical problems by semantic ascent. This turn gained us some important insights at the cost of a fundamental misdirection of philosophical energy. The appeal to sets was one effect of the immense development of set-theory in this century. This raised the hope of resolving all sorts of philosophical problems by a series of set-theoretical technical fixes.

To appreciate the utter implausibility of the attempt to evade properties by means of predicates it is perhaps sufficient to consider a case where a thing's property changes. A cold thing becomes hot. For one who puts his or her faith in predicates this is a matter first of the predicate 'cold', or its semantic equivalent, *applying* to or *being true of* the object, and, second, the predicate 'hot' becoming applicable after 'cold' loses applicability. Properties in the object are but metaphysical shadows cast on that object by the predicates.

But what have predicates to do with the temperature of the object? The change in the object could have occurred even if the predicates had never existed. Furthermore, the change is something intrinsic to the object, and has nothing to do with the way the object stands to language.

I think that one has to be pretty far gone in what might be called Linguistic Idealism to find predicates much of a substitute for properties. But sets are a somewhat more serious matter. After all, to substitute classes for genuine properties is at least to remain a realist, even if a reductive realist, about properties. Even so, an account of properties in terms of classes is still full of difficulties.

First, there is what might be called the 'Promiscuity problem' – a fairly close relative of the grue problem. Sets abound, and only a very few of them are of the slightest interest. Most of the uninteresting ones are uninteresting because they are utterly heterogeneous, that is, the members of the set have nothing in common. In particular, for most sets there is no common property, F, such that the set is the set of *all the Fs*. The result is that mere sets are insufficient to given an account of properties: at best having a property is a matter of membership of a set *of a certain sort*.

Indeed, not only are most sets too poor to support properties, others,

it seems are too prosperous, and support more than one property. This is the problem for a class account of properties that all philosophers are conversant with. It is the coextension problem, the problem of the renates and the cordates, the creatures with kidneys and the creatures with hearts.

Returning to the Promiscuity problem, which I judge to be a much more serious fundamental problem, there are various ways that an account of properties in terms of class might move under pressure. One solution, pioneered by Anthony Quinton (1957, 1973), is to introduce a new, fundamental, and so not further analyzable, notion of a *natural* class. Some classes are natural, most are not. The natural ones admit of degrees of naturaliness, but no analysis of naturalism is possible.

Of the difficulties that such an account faces, I shall here call attention to but one. (A problem concerning relations will be mentioned when the resemblance theory is discussed.) It is similar to the difficulty urged a moment ago against an account of properties in terms of predicates. It was said that when a thing changes temperature, it is the thing itself that changes. The change in the applicability of certain predicates is, fairly obviously, subsequent and secondary. In the same way, consider the natural class consisting of all and only the objects having temperature T. Let *a* be a member of this class. What have the *other* members of the class, or at any rate the other members that are wholly distinct from *a*, to do with *a*'s temperature? After all there would appear to be a possible world in which these other members do not exist, or where they exist but lack temperature T.

Somewhat more attractive than a Natural Class theory is a Resemblance account. According to one version of this view, talk about properties of a particular has as its ontological ground a suitable relation of resemblance holding between the particular in question and suitable paradigms. It might seem that such a view falls victim to the argument just advanced against Predicate and Class accounts. What have the paradigms to do with the *being* of the properties of things that suitably resemble the paradigms? I used to think that this was a good argument against a Resemblance analysis as well as Predicate and Class accounts. But I have recently come to think that the consideration that resemblance is an *internal* relation, based upon the nature of its terms, will block the argument in the case of a Resemblance theory (see my 1989, Ch. 3, Sec. II). Details must be left aside here.

But it is worth noticing that the Resemblance theory, like a Class theory

(and a Natural Class theory), is unable to distinguish between different but coextensive properties. In a paradigm version, for instance, it would not be possible to set up different paradigms for the different properties. In any case, the detail required to work out a Resemblance theory is considerable, and trouble may lurk in the elaborate constructions required. There is also trouble concerning relations. The problem is that when *a* has R to *b*, and *c* has 'the same' *R* to *d*, the resemblance has to hold between the way *a* stands to *b*, or the one hand, and the way *c* stands to *d*, or the other. This formulation already involves the notion of relation in the phrase 'stands to'. How to eliminate this? It seems that the Resemblance theory will have to use the same device that a Class theory uses, and identify the terms that resemble with the ordered sets $<a,\ b>$ and $<c,\ d>$. This still involves the relational notion of *order*, and if that is to be eliminated the device of Wiener or Kuratowski will have to be employed and each ordered pair identified with unordered classes of classes. This has a consequence that is also a consequence for a class theory: different classes of classes will each serve as *a*'s having *R* to *b*, and, much worse, the same class of classes can be used for different relations between *a* and *b*. Such arbitrariness strongly suggests that the classes in question do no more than represent, map, the state of affairs of *a*'s having *R* to *b*. The classes are not *identical* with the state of affairs, which is what is needed for metaphysical analysis.

A final criticism that I will make of the Resemblance theory leads us directly to the postulation of properties. I begin by noticing that a traditional argument against the Resemblance analysis is that the resemblance relation is not a two-termed but a three-termed affair. If *a* resembles *b*, in general, they resemble in certain *respects*, and fail to resemble in other *respects*. But respects are uncomfortably close to properties, which the Resemblance theory proposes to do without.

I do not think that this traditional objection is at all conclusive as it stands. The Resemblance theorist can argue that the metaphysically fundamental relation of resemblance is two-termed (though admitting of degrees like the relations of *being distant from* or *more massive than*). It can then be argued that respects and resemblance in respects supervene upon the network of two-term resemblances which are fundamental. But the Resemblance theorist remains in some embarrassment when he comes to explain the formal properties of his fundamental relation. He has to say that the two-termed relation is non-transitive. There is an exception: the limiting case of exact resemblance. But in general: if *a* resembles *b* to

degree D and b resembles c to the same degree, the degree to which a resembles c can take any value. Why is this? The Resemblance theory, it seems, must take it as a primitive, not to be further analyzed, fact. A Property theory, however, can *derive* this non-transitivity. It is a matter of a resembling b in respect of a certain set of properties, b resembling c in respect of a *different* set of properties. This can naturally be expected to produce a different degree of resemblance between the pairs $\langle a, b \rangle$ and $\langle a, c \rangle$. The transitivity of exact resemblance is also explained, since in such a case the properties of a, b and c are the same. Explanatory power is a virtue, and lack of explanation a defect, in metaphysics as much as science.

The above argument led us from resemblance to properties. But I believe that the explanatory power of a theory which gives real existence to properties (and relations) is seen most clearly in connection with *causation* and *natural* law. Suppose that the water in the kettle is heated by the fire. We surely want to deny that it is the whole fire, qua token, that causes the heating of the water. The fire, first of all, must be in the right *relations* to the kettle, say underneath, and the kettle must in turn *contain* the water. Still more importantly, the fire must be *hot*. Consider how this is explained by an account in terms of predicates. The predicate 'underneath' applies to the pair of the fire and the kettle, the predicate 'hot' to the fire and, eventually, to the water. But when we have said that these predicates apply, we have surely not said enough. The situation cries out for explanation. It is surely something definite *about* these three things that allows the predicates to apply. Must there not be something quite specific about the things which allows, indeed ensures, that these predicates apply? The predicates require *ontological correlates*. The predicate theory does have correlates indeed, but they are no more than the objects themselves, and so are far too coarse.

It is little better to appeal to classes, even natural classes. What has this fire's heating this water in this kettle, here and now, to do with the fire's membership of the class of hot things? A satisfactory correlate must be found 'within' the fire. A sophisticated Resemblance theory can, I think, appeal to the *natures* of the resembling things, natures from which the resemblances flow. The natures are suitably internal, but are as coarse as the things themselves, (indeed, should perhaps be identified with the things themselves).

As with causes, so with laws. I am not speaking of law-*statements* but of the ontological correlates of true law-statements, that in the world

which makes true law-statements true. Suppose it be a truth that gravitational force between bodies is equal to the product of their masses divided by the square of their distance. This appears to be a certain connection between the properties of massive things. One can try to translate the corresponding law-statements into statements of universal quantification where the subject-terms are nothing but first-order particulars. But although statements about first-order particulars may follow from law-statements, the latter as is well known, say something more, a more that is plausibly a link between properties. And even if one denies this, perhaps because one thinks that properties are not universals but particulars, it still seems that the ontological correlates of true law-statements must involve properties. How else can one pick out the uniformities which the law-statements entail?

Why has there been such hostility to properties (and relations) among so many contempory leaders of analytic philosophy? Is it just the Ockhamist spirit? Do without properties and relations if you can! Or it is the influence of Quine, with his doctrine that the predicate of a true statement carries no ontological implications? (Together with his nasty remarks about 'McX', the upholder of universals – Quine, 1961.) All these things, maybe, and others. But I think that upholders of properties and relations also have something to answer for. As so often happens, in philosophy and elsewhere, an excellent case has been ignored because its advocates overdid things and made exaggerated claims. Simply put, they found far too many properties.

What has happened is that for each distinct predicate, upholders of properties have been inclined to postulate, corresponding to it, a distinct property. Synonymous predicates, 'father' and 'male parent', were generally thought to apply in virtue of just one property. But for non-synonymous predicates, each its own property. To self-contradictory predicates, perhaps, no property corresponds. But for each of the rest, a property of its own.

As a very beginning one might eliminate from this monstrous regiment of properties all those where the corresponding predicate fails to apply, is not true of, anything. After all, the argument for properties that I advanced was for something in particulars which would allow the application of predicates. No application, no property. There is a tendency, whose rationale I do not really understand, to think of properties as necessary beings. A necessary being, if it is possible, exists, and so, if properties are necessary beings, all non-self-contradictory

properties exist. But if, as I think we should, we take properties to be contingent beings, then it seems reasonable to deny that there are uninstantiated properties.

This is not to say that it may be convenient from time to time to talk about, to make ostensible reference to, uninstantiated properties. No body is perfectly elastic, so there is no property of *being perfectly elastic*. But it may be useful to compare more or less elastic bodies in the degree to which they approach the unreachable perfect elasticity. A useful fiction, however, is still a fiction.

If predicates actually apply to, are actually true of, things, then, of course, it is perfectly legitimate to introduce *a sense* in which the things automatically have a property corresponding to just that predicate. Indeed, this is a very useful sense, a point that I have in the past tended to overlook. To make use of Carnap's phrases, the *material mode* is much less fatiguing to the imagination and the intellect, than is the *formal mode*. Such properties, however, cut no ontological ice. The properties that are of ontological interest and which we are concerned with here, are those constituents of objects, of particulars, which serve as the ground in the objects for the application of predicates. And concerning these properties, the true properties I am inclined to say, there is no reason to think that to each distinct predicate that has application corresponds its own distinct property in the object. Indeed, there is much reason to think the opposite.

Instead of approaching the matter of such properties directly, it may be helpful to think in the first place in terms of 'good' and 'bad' predicates, where good and bad are to be assessed in terms of our purely theoretical interests: the sort of predicates that the spectator of all time and eternity might find attractive. And here, I think, we are led on to Plato's marvellous image of carving the beast (the great beast of reality) at the joints. The carving may be more or less precise. But as the carving is the more and more precise, so we reach predicates that are of greater and greater theoretical value, predicates more and more fit to appear in the formulations of an exact science. At the limit, monadic predicates apply in virtue of strict properties. An upholder of universals will conceive of these properties as strictly identical in their different instances. A believer in particularized properties, in tropes, will deny identity but allow the symmetrical and transitive relation of exact similarity. It is properties thus conceived that serve as the ontological foundation for the application of predicates, most predicates at any rate, to first-order particulars.

How do we determine what these ontological properties are? The answer, in part, is the same as the answer to the question 'How do kangaroos make love?' With difficulty. In the epistemic state of nature, the only predicates to which we can give much trust are those suggested by ordinary experience and ordinary life. We cannot but take it that these predicates carve out properties that approximate to some of the joints to some extent. In that state of nature, we cannot but think that blue is better than grue. But in the present age we take ourselves to have advanced beyond the epistemic state of nature, and to have sciences that we speak of as 'mature'. There we will find the predicates that constitute our most educated guess about what are the true properties and relations. Property-realism, whether the properties be taken as universals or particulars, should be an *a posteriori*, a scientific, realism.

If we combine an *a posteriori* or scientific realism about properties (and relations) with the speculative but attractive thesis of physicalism, then we shall look to physics, the most mature science of all, for *our best predicates so far*. Physics (perhaps it will have to be a cosmological physics as well as the physics of the very small) shows promise of giving an explanatory account of the workings of the whole space-time realm, and thus, perhaps, the whole of being. And it shows promise of doing this in terms of a quite restricted range of fundamental properties and relations. These properties and relations are for the most part quantitative in nature, and the laws that govern them are functional in nature. I will just note that quantities and functions seem to me to involve rather deep problems for the property-realist. (Happily, though, the problems for the alternative positions, such as Predicate and Class Nominalism, seem to be far worse.)

Keith Campbell has suggested, in his new book *Abstract Particulars* (Blackwell, 1990, p. 13), which puts forward a trope metaphysics, that a metaphysics of this physicalist sort is not particularly economical with properties. For suppose that some fundamental quantity such as length is really continuous. We will then be faced with the necessity to postulate continuum-many length-properties corresponding to each different length taken as a type. Some lengths may not be instantiated, but that will not bring the number down.

Continuum-many properties is a lot of properties, to be sure. But let us remember a remark that Mr. Reagan made when he was Governor of California. Speaking of the Sequoia tree, he said 'seen one, seen them all.' If you have seen one length, then given only some mathematics, which is

topic-neutral, you can grasp the notion of lengths of any length. The class of length-types is a unitary thing, and in taking lengths to be fundamental properties, if you do so, you are making a quite economical postulation. And it may be that a relatively small number of quantities such as length are the only fundamental quantities that physics requires us to postulate.

2. UNIVERSALS VS. TROPES

So much in defense of properties, although much more could be said. In the second part of my paper I will take up an issue *within* the theory of properties, an issue that has enjoyed quite a lot of recent discussion. It is the question whether we should take properties to be universals or particulars. There are those who admit both universal and particular properties into their ontology. Perhaps Aristotle and even Plato were among them. But I think that this position sins against economy. If you have universals, you can do without the particularized properties, and *vice-versa*. So for me, and I think for most contemporary metaphysicians, the question is which should we choose.

I was brought up by my teacher, John Anderson, to reject the Particularist position. (He used to criticize G.F. Stout's view.) I still favor the Universalist view, but recently I come to think that tropes have more to be said for them than I have allowed previously. In particular, I now see more clearly how tropes can serve as substitutes for universals in many respects.

A trope theory is best combined with a resemblance theory, and developed as a sophisticated Resemblance Nominalism. Of particular importance here is the notion of *exact resemblance*. If we work with ordinary particulars, then, with the possible exception of such things as fundamental particles, exact resemblance is a theoretical ideal. We all remember Leibniz's unfortunate courtiers searching vainly for identical leaves in the garden. But if we move to the much thinner *tropes*, then exact resemblance may be achievable. Two tropes that are constituents of different things might resemble exactly in mass, in length, in charge, and so on. The plausible examples are again found at fundamental levels. Thus, it is at present believed that the charge on each electron is exactly the same. 'Exactly the same' appears to assume *identity* of charge in different electrons. But it can be rendered in the language of tropes by saying instead that the different charge-tropes associated with the different electrons are all exactly similar. The interesting thing about exact

similarity, is that is symmetrical *and transitive*. (Less than exact similarly is not transitive, even for tropes.) As a result, the relation of exact similarity is an equivalence relation, partitioning the field of tropes into equivalence classes. Tropes will then do much the same work as universals. Suppose that a believer in universals and a believer in tropes have co-ordinated their views in the following way. For each universal property postulated, the trope theorist postulates a class of exactly similar tropes, with universals and tropes properties of the very same class of things. For each class of exactly similar tropes postulated, the Realist postulates a class of thing which all have the same property, with tropes and universals properties of the very same class of things. What inferiority is there in the trope theory?

I used to think that the Universals theory had an important advantage here. Where we have what the trope theorist thinks of as exact similarity of tropes, we do not scruple to speak of *sameness* of property. Even a trope theorist will allow that by the lights of our present physists electrons have *exactly the same* charge. But 'same' means identical does it not? Yet the trope theory denies identity.

However, I have become convinced that in our ordinary usage 'same' does not always mean identical. There is what Bishop Butler so brilliantly characterized as a 'loose and popular' sense of the word 'same' (1736). Butler was thinking about the replacement over time of particles in an object such as a human body. We say the *same* body but we don't really mean it. It is only a loose and popular identity. By itself, even if we accept it, Butler's point is rather frustrating. What rules are we going by when we use 'same' in the loose and popular way? Here I am indebted to a Sydney student, Peter Anstey. He suggested that we are prepared to use 'same' in this relaxed way only if the things said to be the same are both members of a relevant equivalence class. Though different, the things said to be the same must both be members of the same class, where 'same class' is, of course, taken in the *strict* sense.

If one takes *portions of the lives of organisms* as a field, then it seems that they fall into equivalence classes, where the members of any one class constitute the totality of the life of a single organism (fission, fusion and so on being neglected). It is of course difficult to spell out just what the equivalence relation is: 'identity over time' is a puzzling subject. But, if Anstey is right, it must be in virtue of this equivalence relation that we assert 'identity', and assert it even though we believe that *strict* identity is not involved. (A further suggestion by Anstey. Is this relevant to the topic

of 'relative identity'? When a and b are 'the same F, but not 'the same G', is this because the identity is loose and popular, and two different equivalence relations are involved?)

This is, alas, good news for the tropes. When we say that two electrons have the very same charge, then according to the trope theory *strictly* the tropes involved are not identical. But the two tropes are both members of a relevant equivalence class, where the equivalence relation is exact similarity, and so can be said to be 'the same' in a loose and popular sense.

Unfortunately, this is not all the good news for the Trope theory. A very important topic in the theory of properties (and relations) is that of their *resemblance*. Particulars resemble: that is clear enough. But so do properties. The colors all resemble each other, so do the shapes, the masses, the lengths. One property can resemble another more than it does a third. Redness is more like orange than it is like yellow. A kilo is more like a pound than it is like an ounce.

We may think of the whole field of properties as arranged in a multidimensional order. This order appears to be largely objective, and is to be interpreted as a resemblance-order. For properties to stand near to each other in the order is for them to resemble each other quite closely.

If these properties are universals, then it will be natural to construe these resemblances between properties in accordance with the old slogan 'all resemblance is partial identity'. That is how I construe it myself. Resembling universals have common constituents, with either one of the properties containing all the constituents of the other universal and more besides, or else a mere overlap in constituents. I say 'constituents' rather than 'part' because I think that this partial identity is not the simple sort of partial identity envisaged by the mereological calculus, the calculus of whole and part. (A point that confused me for many years.) I cannot go further into the matter here. To do so would involve getting into a huge new topic: the theory of facts or states of affairs. (See my 1989.)

But however all this may be, an upholder of tropes can deal with the resemblance of properties in a way that parallels the treatment of the topic by an upholder of universals. We have seen the potential to set up a one-one correlation between properties taken as universals, on the one side, and equivalence classes of exactly similar tropes, or the other. To make the matter vivid, select just one trope from each of these equivalence classes and range each of these tropes opposite to its corresponding universal. This structure of tropes will exactly reflect the multi-dimensional resemblance-structure of the universals.

How resemblance is interpreted will presumably differ in the two structures. The trope theory is not under pressure to interpret resemblance between tropes as partial identity, a move that is indeed against the spirit of trope theory, although that option would be open. (Similarly, it is an option for the universals theory to treat resemblance between *universals* as primitive and unanalyzable, although that goes against the spirit of a universals theory.) A trope theory, with exact resemblance already treated as a primitive, will presumably embrace the view that, in fundamental cases at least, lesser degrees of resemblance between tropes are also primitive and unanalyzable. But the point I want principally to make here is that the Trope theory is in as good a position as the Universals theory to deal with the difficult topic of resemblance of properties. The friends of the tropes can say to the friends of the universals: 'Anything you can do, I can do better, or at least equally well'.

But I finish now by saying I do not believe in the tropes. First, there is the question, already touched upon, of the Axioms of Resemblance. The trope theorist requires such axioms. *First*, there is symmetry. If *a* resembles *b* to degree *D*, then *b* resembles *a* to degree *D*. *Second*, there is failure of transitivity. If *a* resembles *b* to degree *D*, and *b* resembles *c* to degree *D*, then it is not normally the case that *a* resembles *c* to degree *D*. This holds for tropes as much as for ordinary particulars. However, *third*, transitivity is restored for a special case. If *a* exactly resembles *b*, and *b* exactly resembles *c*, then *a* exactly resembles *c*. This transitivity is a particular case of a more general principle: if *a* resembles *b* to degree *D* and *b* exactly resembles *c*, then *a* resembles *c* to degree *D*. Resemblance to degree *D* is preserved under the substitution of exact resemblers.

For the Trope theorist these necessities are *brute* necessities, fundamental necessities that cannot be explained further. The Universals theory need carry no such ontological baggage. The symmetry of resemblance is simply the symmetry of identity. The transitivity of exact resemblance is the transitivity of identity. The non-transitivity of ordinary resemblance is the non-transitivity of partial identity. The Axioms of Resemblance are but particular cases of the axioms that govern identity.

Explanation is a virtue in metaphysics, as elsewhere. I submit that this startlingly easy deduction of the properties of resemblance from the entirely uncontroversial properties of identity is a major advantage of the Universals theory. It enables one to see the intuitive force behind the old, inconclusive, criticism brought against Resemblance Nominalisms that resemblance is resemblance *in identical respects*.

My second reason for rejecting the Trope theory is more controversial, depending as it does on views that would be contested by many. It is that I think that universals are required to get a satisfactory account of laws of nature.

I note again that by laws of nature I mean not true law statements, but that entity, state of affairs, in the world that makes true law statements true. I believe that the contemporary orthodoxy on laws of nature – that basically they are mere regularities in the four-dimensional scenery – is in a similar position to that enjoyed by the regimes in power in Eastern Europe until a few months ago, if 'enjoyed' is the right word. No doubt the end to Regularity theories of law will not come so suddenly, though, because inside their own subject philosophers are great conservatives.

Regularity theories of laws face the grue problem. That, I think, can only be got over by introducing properties, sparse properties, into one's ontology. However, the properties could, I think, be tropes as well as being universals, so there is no advantage to universals here. More to the present point, even with properties given, Regularity theories make laws into *molecular* states of affairs. These tokens of a certain phenomenon behave in a certain way, so do these, so do all instances of the phenomenon. There is here no *inner connection* between, say, cause and effect in the individual tokens that fall under the causal law. This conclusion can, I think, be enforced by noting with Reichenbach and others that only some cosmic regularities are manifestions of law; by the difficulty in seeing how such a molecular state of affairs could 'sustain counterfactuals'; and by the incredible shifts that are necessary to accommodate probabilistic laws within a regularity approach.

Only a higher-order fact about the universals involved in the individual positive instances falling under the law can, as far as I can see, provide the atomic state of affairs that will solve these difficulties. If *being an F* ensures or makes probable to some degree that the F, or something related to it, is a G, with F and G universals, then I think that an internal connection is provided. More controversially, I think it can also be seen that such a connection automatically, analytically, and yet non-trivially, provides for a regularity or statistical distribution to flow from the connection. Indeed, I think that, although postulating such a connection does not cure wooden legs or halt tooth decay, it does go a great way to help us with the problem of induction. (For all this see Armstrong, 1983.)

So: my idea is that a Universals theory can provide us with a satisfactory

account of laws of nature, while a Trope theory cannot. It is a controversial and complex argument, which cannot be assessed in any hurry. But even without this, the Trope theory still needs its Axioms of Resemblance, and that is a clear-cut disadvantage. I know of no such compensating disadvantage for the properties are universals.

REFERENCES

Armstrong, D.M. 1983 *What Is A Law of Nature?*, Cambridge University Press, Cambridge, England

______, 1989 *Universals, An Opinionated Introduction*, Westview Press, Boulder, Colorado

Butler, J. 1736 'Of Personal Identity' in *The Analogy of Religion*

Campbell, Keith 1990 *Abstract Particulars*, Blackwell, Oxford

Quine, W.V. 1961 'On What There Is' in *From A Logical Point of View*, Harvard University Press, Cambridge, MA

Quinton, A. 1957 'Properties and Classes,' *Proceedings of the Aristotelian Society*, 58

______ 1973 *The Nature of Things*, Routledge and Kegan Paul, London

Skyrms, B. 1981 'Tractarian Nominalism', *Philosophical Studies*, 40.

UWE MEIXNER

ON NEGATIVE AND DISJUNCTIVE PROPERTIES

In *Universals and Scientific Realism*, II, D.M. Armstrong argues against the existence of negative and disjunctive universals, in particular against the existence of negative and disjunctive properties – monadic universals. He admits, however, conjunctive universals. My paper is a defense of negative and disjunctive properties.

By the word "term" I shall always mean *singular term*; property-terms are special singular terms; *prima facie* they refer to properties. In the present context $\wedge$, $\neg$, $\vee$ are term-forming operators; they form more complex property-terms out of simpler property-terms; for example, if f and g are property-terms, $(f \wedge g)$, -f, and $(f \vee g)$. These complex property-terms can be easily eliminated in certain contexts. We have:

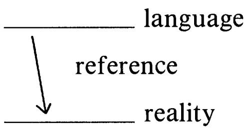
$(f \wedge g)(x)$ iff $f(x)$ and $g(x)$ [$f(x)$: x exemplifies f]
$(f \vee g)(x)$ iff $f(x)$ or $g(x)$
$(\neg f)(x)$ iff not $f(x)$

These rules suffice to make any occurrence of a conjunctive, disjunctive or negative property-term disappear in *contexts of first-order predication* in favor of simple property-terms. So far there is no need to take complex property-terms ontologically seriously. Unfortunately contexts of first-order predication are not the only contexts in which complex property-terms occur. Take for example "Not to lie is sometimes hard", "He intends to fly or to go by train". And in deference to Armstrong's distrust of the ontological authority of ordinary language I offer another example. Let "f" and "g" refer to two impeccable monadic universals, which play a central role in the higher regions of quantum mechanics; let us stipulate that there are some particulars to which they both apply, and some particulars to which only one or the other applies; these universals

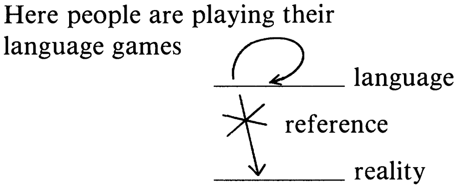
will serve as our examples throughout this paper. Then we can formulate the sentences "$\neg$ f is not a property", "(f $\vee$ g) is not a property".

Since "$\neg$ f", "(f $\wedge$ g)", and "(f$\vee$g)" will be seen to be definable by definite descriptions, there is a general procedure for eliminating them from contexts like those mentioned; take Carnap's or Russell's method. Nevertheless we shall see that they are *genuinely referring* singular terms. The singular term "the father of John" is also eliminable from any context; nevertheless it is a genuinely referring singular term: since there is exactly one father of John, it genuinely refers to John's father. "$\neg$ f", "(f $\wedge$ g)", "(f $\vee$ g)" will analogously be seen to genuinely refer to something, that is, *not* to some entity artificially stipulated as their reference. The only question is whether this something is indeed a property.

Let me make a parenthetical remark. Abstract singular terms in subject-position are the stronghold of the Platonists against the nominalists. This stronghold is unconquerable as long as both parties assume, as I do here, the *reference-paradigm* of language and its corresponding concept of truth:



Nominalists had better give up this paradigm. Yet many nominalists will find it hard to swallow the relativist consequences of giving it up in favor of the non-referential paradigm advocated, for example, by Wittgenstein in the *Philosophische Untersuchungen*; this paradigm might be sketched like this:



Moreover, the referential paradigm of language is very well entrenched. Motives of ontological parsimony aside, there is hardly any reason to relinquish it.

I proceed on the assumption, which is to be justified later, that "$(f \wedge g)$", "$(f \vee g)$", and "$\neg f$" genuinely refer to something.

Armstrong says that there are no negative and disjunctive properties. This implies that he considers the sentences "$\neg f$ is not a property", "$(f \vee g)$ is not a property" to be true in their straightforward meaning. If he did not consider them true in their straightforward meaning, I, for my part, would not know what he means by negative and disjunctive properties. I assume he means this:

h is a negative property	:=	h is a property and $\exists g'$(g' is a property and $h = \neg g'$) ($\exists$: there is)
h is a disjunctive property	:=	h is a property and $\exists g' \exists f'$(g' is a property and f' is a property and $h = (f' \vee g')$)

If $\neg f$ and $(f \vee g)$ are not properties, what are they? Of course we can stipulate that the negation of a property and the disjunction of two properties never is a property, but is rather a so-and-so (a mere *concept*, for example); but why should we make such a stipulation? Why not admit $\neg f$ and $(f \vee g)$ as properties along with f and g?

Armstrong offers several arguments. I cannot go over all of them here. I shall restrict my attention to two of them, which seem to me to be central. Armstrong says: "Properties should be such that it at least makes sense to attribute causal powers to objects in virtue of these properties. But how could a mere lack or absence endow anything with causal powers?" (*Universals and Scientific Realism*, II, p. 25). This is directed against negative properties; no negation of a property is a property, since it is "a mere lack or absence", which cannot confer causal powers, something every property ought to be able to do. I will not at this point impugn the conception of properties implied. Just one global remark: the concepts of causality and privation are in my eyes heavily epistemological and evaluative in character, and hence out of place in ontology. But the negation of property f *is not* "a mere lack or absence", whatever that may be. For then the negation of this negation would have to be a mere lack or absence, too. "Nothing will come of nothing", as Armstrong says. But the negation of the negation of f *is not* a mere lack or absence; rather it is the property f itself: $\neg\neg f = f$.

Armstrong also says: "disjunctive properties offend against the principle that a genuine property is identical in its different particulars"

(*Universals and Scientific Realism*, II, p. 20). I do not impugn the conception of properties implied. But do disjunctive properties really offend against the mentioned principle? What, indeed, have two particulars in common that both exemplify $(f \vee g)$, because one of them exemplifies f, but not g, the other g, but not f? What is identical in them because of this? This seems a hard question. But in fact there is no problem at all. They have in common the greatest entity that is both a part of f and a part of g. (Armstrong, it will be remembered, allows us to speak of parts of properties.) This entity shared by them is no other than $(f \vee g)$ itself. Assume a first-order language (variables f, g, h, ...) including a part-predicate $\leq$ (sentences $f \leq g$, ...); call the entities in its domain "concepts"; let us *at this point* assume that some, but not all, concepts are properties.[1] Then we can define:

$(f \vee g) : = \iota\, k[k \geq f \text{ a. } k \leq g \text{ a. } \forall m(m \leq f \text{ a. } m \leq g \text{ imp. } m \leq k)]$
(a. : and; imp. : implies; $\forall$: for all; ι : the)

Now, it seems very plausible that every part of a property is a property. Armstrong himself admits this; on page 35 of *Universals and Scientific Realism*, he writes: "a part of a property will be a property". Hence, since $(f \vee g)$ is a part of property f (and of property g), it is a property itself. Hence there are disjunctive properties.

For comparison I also offer the definitions of conjunction and negation (of concepts):

$(f \vee g) := \iota\, k[f \leq k \text{ a. } g \leq k \text{ a. } \forall m(f \leq m \text{ a. } g \leq m \text{ imp. } k \leq m)]$
(The conjunction of f and g is the smallest concept of which both f and g are parts.)

$\neg f := \iota\, k[\forall h(h \leq k\text{a. not } \forall h'(h \leq h') \text{ imp. not } h \leq f)\text{a. } \forall g'$
$(\forall h(h \leq g'\text{a. not } \forall h' (h \leq h') \text{ imp. not } h \leq f) \text{ imp. } g' \leq k)]$
(The negation of f is the greatest concept that has no non-trivial part in common with f.)

Where there are parts, there are sums, products, and complements. The sums, products, and complements of individuals are individuals. Why shouldn't the sums, products, and complements of properties be properties? From the mereological point of view Armstrong offers no good reason why $\neg f$ and $(f \vee g)$ are not properties, but $(f \wedge g)$ is.

It should be noticed that for the purposes of this defense it is not

necessary to assume that every property has a complement, that every two properties have a product. As we know, this is not true of individuals. I am inclined to say that it is true of concepts, and hence of properties (since every property is a concept). But be that as it may, what good reason is there for denying that *some* property has a complement, or for denying that *some* pair of properties has a product? None. Rather, there is good reason for accepting the points in question, if we want to speak of parts of properties at all and apply the common laws of mereology. Since this is so, and since complement and product are unique if they exist, we may unproblematically assume that f has exactly one complement, referred to by "$\neg$ f", and f and g exactly one product, referred to by "(f $\vee$ g)". And why shouldn't these also be properties? Clearly, that they are *not* properties cannot be seen from their definitions.

But assume that $\neg$ f and (f $\vee$ g) are not properties. Strangely enough $\neg\neg$ f *is* a property, since $\neg\neg$ f = f and f is a property; this is like saying that (2:0) is not a number, but (2:0):0 is. This is not an important point. But the next one is. Let us for a moment assume that g is a part of f; then we have (f $\vee$ g) = g; hence (f $\vee$ g) is a property, since g is a property. This means that if Armstrong wants to uphold the position that every disjunction of two properties is not a property, he must also hold that no property g ever is part of another property f. But the latter position seems to be clearly false, and contradicts moreover Armstrong's position on conjunctive properties. If there are conjunctive properties, then, of course, they have properties as parts.

Given Armstrong's criteria for propertyhood, like causal efficacy and identity in different particulars, can we be a priori sure that in every case, if h and k fulfill these criteria, $\neg$ h and (h $\vee$ k) don't? Why shouldn't it be possible that besides h and k also $\neg$ h and (h $\vee$ k) are causally efficacious and identical in different particulars? Let me give three examples, that are not mere possibilities.

(1) *Even* and *not-even*, that is, *odd* partition the domain of natural numbers (the entities they can be a predicated of) into two halves of equal size. Even numbers are the values of the function 2x, odd numbers are the values of the function 2x + 1 (with respect to the domain of natural numbers). Can we say that being odd is a mere absence or privation? We cannot. Rather, oddness is a substantial quality present in all odd numbers, just as evenness is a substantial quality present in all even numbers.

(2) *Animate* and *inanimate* partition the domain of material objects (the entities they can be predicated of) into two halves. Can we say that being inanimate is a mere absence or privation? We cannot. It certainly is a substantial quality just as being animate is. This is shown in the fact that only inanimate objects can serve as nourishment for living beings:

hence there is a causal efficacy that only inanimate objects can exert. To die is not a mere event of privation, but also a release of forces so far locked up inside the organism, forces that will serve to uphold life in other organisms.

It is not possible to explain the specific causal efficacy of inanimate objects ultimately by their so-called positive properties, since only being inanimate causes the object to have these positive properties. Armstrong, however, thinks that prima facie "negative" causation can always be seen to be "positive" causation, for example in the case of death by dehydration. I quote: "To say that lack of water caused his death reflects not a metaphysics of the causal efficacy of absences but merely ignorance. Certain (positive) processes were going on in his body, processes which, in the absence of water, resulted in a physiological condition in virtue of which the predicate 'dead' applied to his body" (*Universals and Scientific Realism*, II, p. 44). What is revealing about this passage is that the author finds it necessary to insert into his purported positive account of death by dehydration the phrase "in the absence of water". Apparently even the author darkly feels that absence of water caused the positive processes he invokes to occur; they could not have occurred without it; it, and not they, is ultimately responsible for death by dehydration.

(3) Finally, *male* and *female* have a greatest common part: *male-or-female*, which is identical with *sexed*. Clearly, there is a causal efficacy that only sexed objects can exert, there is a common quality in all sexed objects.

For those not prejudiced in favor of the concepts of advanced natural science these examples serve to show that besides a property h and a property k we can have a property $\neg h$ and a property $(h \vee k)$ (on Armstrong's own criteria for propertyhood!), hence that there are negative and disjunctive properties. For others they pose a question: Might not what they exemplify also occur in the domain of quantum mechanics?

The criticism of Armstrong's position I am now going to offer is more fundamental than anything said so far. Armstrong thinks the semantical tradition in ontology to be on the whole discreditable. I am not of this opinion. But be that as it may, there is a profound reason why the semantical tradition has reigned in ontology and – we may safely assume – will continue to reign. Ontology will not come to lean on natural science for the reason that it seeks exclusively *metaphysically necessary*[2] truths, which are not germane to natural science, but rather to semantics. (In what follows I take the word "necessary" to mean *metaphysically necessary*.) A classical task of ontology is to classify entities on the most general level, that is, to decide the truth of questions of the form "Is b a C?", where C is any ontological category, like "universal", "relation", "individual", "state of affairs", and so on. Hence the following principle must be true:

(P) For every ontological category C: "b is a C" is necessarily true or "b is not a C" is necessarily true

If it were not true, then for some ontological category K "Is b a K?" would not be an ontological question, since the correct answer to it, whether positive or negative, would not be necessarily true, and ontology exclusively seeks answers that are necessarily true, as I have said. But it is a task of ontology to decide truthfully *any* question of the form "Is b a C?", where C is any ontological category. Consequently "Is b a K?" turns out to be an ontological question after all, and we have a reductio argument for the truth of (P).

From (P) (P') follows:

(P') "b is a property" is necessarily true or "b is not a property" is necessarily true,

if "property" is an ontological category, that is, without epistemological or evaluative connotations, which still seems to me to be the most satisfactory view of the meaning of that word. From (P') we can see that natural science is irrelevant for the question whether an entity y is a property or not. The truths that belong to natural science are all metaphysically contingent; but the question at hand has an answer that is necessarily true. If we already know that an entity y is a property, then natural science may tell us contingent facts about it: that it is exemplified, that it plays a fundamental role in the make-up of the world, while its negation does not; that it lends a certain causal efficacy to the particulars that exemplify it, while its negation does not. All this is valuable information, but it is of no concern for ontology. For, from the metaphysical point of view, all this might be different, y might not be exemplified; or it might not play a fundamental role in the make-up of the world, but at the same time its negation might; it might lend no causal efficacy to the particulars that exemplify it, but at the same time its negation might lend causal efficacy to the particulars that exemplify it in their turn. And yet the entity y would remain a property, since it is a property by metaphysical necessity. Its being a property is independent of what natural science can tell us about it; in particular there is no link between being a property and causality. But rather its being a property ("property" taken in a purely ontological sense) is founded by the role expressions corresponding to it play in our descriptively meaningful discourse, which in order to be effective must mirror not the particular

structure of reality, but the general structure of any possible reality. Being a property is one feature of that general structure, which is the structure ontology seeks to describe; the particular structure of reality is the concern of physics.

Now, expressions corresponding to f and $(f \vee g)$ play the very same role in our descriptively meaningful discourse as do expressions corresponding to f and g. Language is not an infallible guide in ontological matters (by no means!), but it is the best guide we have. Hence f and $(f \vee g)$ are properties along with f and g. Hence there are negative and disjunctive properties.

NOTES

[1] All properties, however, are concepts. If we want to make a distinction between properties and concepts (as I am doing here, but vide infra), then, roughly, properties are concepts which are "important" for *us*; there is, however, no purely ontological distinction between properties and concepts. Concepts can be thought of as the possible intensions (not *meanings*) of one-place predicates. They form ontological families with respect to the entities they can be predicated of. (Being inanimate, for example, does not belong to the family of concepts that can be predicated of natural numbers, but rather to the family of concepts that can be predicated of material objects.) $\leq$, applied to concepts of a certain family, stands for the intensional part-relation between them; being an animal, for example, is an intensional part of being human. This relation (with respect to a domain consisting of the concepts of a certain family) *can* be taken to satisfy the following axioms:

A1 $f \leq g$ a. $g \leq h$ imp. $f \leq h$

A2 $f \leq f$

A3 $f \leq g$ a. $g \leq f$ imp. $f = g$

A4 $\exists g[\forall f(A[f]$ imp. $f \leq g)$ a. $\forall k(\forall f(A[f]$ imp. $f \leq k)$ imp. $g \leq k)]$

A5 $\forall k(At(k)$ a. $k \leq f$ imp. $k \leq g)$ imp. $f \leq g$
[$At(k) := \forall h(h \leq k$ a. not $\forall h'(h \leq h')$ imp. $h = k)$; k is an atom iff every non trivial part of k is identical with k]

A6 $f \leq UgA[g]$ a. not $\forall g(f \leq g)$ imp. $\exists k[k \leq f$ a. not $\forall g(k \leq g)$ a. $\exists h(k \leq h$a. $A[h])]$
[$UgA[g] = \iota g[\forall f(A[f]$ imp. $f \leq g)$ a. $\forall k(\forall f(A[f]$ imp. $f \leq k)$imp. $g \leq k)]$; the conjunction of all A-concepts]

This formalism (a first-order approximation to a complete and atomistic Boolean algebra) is an extension of the calculus of concepts first developed by Leibniz (see W. Lenzen, "Leibniz und die Boolesche Algebra", *Studia Leibniziana* (1984), 16, p. 187-203). According to it every concept has a unique complement (negation), and every two concepts have a unique product (disjunction). One may define in a Leibnizian vein: $f(g) = \forall h($not $h \leq g$ equ. $\neg h \leq g)$a. $f \leq g$ [equ. : equivalent; g exemplifies f iff g is a maximal-consistent concept and f is part of g], which enables one to prove the predication-rules stated at the beginning of this paper. (All this is treated in detail in my forthcoming book *Axiomatische Ontologie*.)

Moreover, add these axioms:

B1	Prop(f) a. $g \leq f$ imp. Prop(g) [Prop(f) : f is a property]
B2	$\exists$fProp(f)
B3	$\forall$ f(A[f] imp. Prop(f)) imp. Prop(UfA[f])

Then one can prove: Prop(f) imp. Prop((f $\vee$ g)), Prop(f) a. Prop(g) equ. Prop((f $\wedge$ g)).

[2] Metaphysical necessity is in between *analytical necessity and nomic necessity*. A sentence in analytically necessary iff it is true, and its meaning and syntactical form are sufficient for its truth; a sentence is nomically necessary iff it is true, and its logical relation to natural laws is sufficient for its truth; a sentence is metaphysically necessary iff it is true in every possible reality (or "world"). Analytical necessity implies metaphysical necessity, which in its turn implies nomic necessity. Let c be the number of concepts; then "There are precisely c concepts" is an example of a sentence that is metaphysically necessary, but not analytically necessary.

REFERENCE

Armstrong, D.M., 1978. *Universals and Scientific Realism*, Cambridge University Press, Vol. 1: *Nominalism and Realism*; Vol. 2: *A Theory of Universals*.

KEITH LEHRER AND VANN McGEE

PARTICULARS, INDIVIDUAL QUALITIES, AND UNIVERSALS

We intend to develop an account of the relation between particulars and universals. Loosely derived from the work of Thomas Reid,[1] the account will be *empiricist*, in that it has our understanding of general concepts dependent upon our prior acquaintance with particular individuals, and it will be *nominalist*, in that it does not require that universals actually exist.

According to Reid, our acquisition of general concepts proceeds in two stages, *abstraction*, in which we come to conceive of the individual qualities of a thing, and *generalization*, in which we mentally conjoin various individual qualities to form a general concept.[2] By abstraction I conceive of the whiteness of this particular sheet of paper; then I generalize, combining my conception of the whiteness of this piece of paper and of the whitenesses of other things to form the general idea of whiteness. The whiteness of this piece of paper is thought of as something peculiar to it, that is not and could not be had by anything else, even something just like this piece of paper in color.[3]

Our story has three players: particulars, individual concepts, and universals. There are two questions that we need to answer about them. First of all, why do we need individual qualities? It is clear that, in order to understand the sentence "Socrates is wise" we need to have a conception of the individual Socrates and the universal wisdom, in the sense that we need to know what the singular term "Socrates" means and what the general term "wise" means; but why do we need a third thing, Socrates's wisdom? Second, what is the ontological status of our three players? We have a conception of particulars (individual qualities; universals), but do particulars (individual qualities; universals) really exist?

To see why we need to conceive of individual qualities in addition to particulars and universals, we consider the alternatives. If we try to get an account of mental acts of predication that includes only particulars and universals, there are only three possibilities. Either our concepts of particulars are constructed out of our concepts of universals or our concepts of universals and particulars arise independently of each other. We consider the three alternatives in turn.

The first alternative is easily dismissed. There is no way we can construct what is ephemeral and concrete out of what is eternal and abstract. Even if we regard particulars as nothing more than bundles of universals, we have to explain the bundling operation. Why is it that some sets of universals occur bundled together and others do not, and how is it that bundles are generated and corrupted?

The second alternative, that we form our conception of universals out of our conceptions of particulars by sorting the particulars into groups, is more promising, but it is still not satisfactory. The problem is that, while we can indeed come to conceive of universals by sorting individuals into groups, unless we have some account of the principles by which we form the groups, the grouping process will be utterly mysterious. We do not group things at random, we group them in virtue of their qualities. Before we can sort things into classes, we first have to know what the things are like.

One way to appreciate the problem is to imagine a world in which all and only red things are round. If we formed our idea of redness simply by sorting individuals into red things and nonred things, we would have no way to distinguish redness from roundness, so that we would regard it as tautological that all red things were round. But this is surely implausible. Even if all and only red things were round, we could imagine a round thing being squeezed into a square mold without changing color, so we could conceive of redness apart from roundness.

The third alternative, allowing that universals and particulars are conceived independently, gives us the right pieces, but no way to put them together. Just having the separate notions of wisdom and Socrates without having the further notion of wisdom and Socrates joined together in the wisdom of Socrates, we cannot form the judgment that Socrates is wise, for we shall not be able to see that wisdom combines with Socrates in a way that foolishness does not. Simply to say that Socrates instantiates wisdom will not suffice; it merely postpones the problem. Why does the ordered pair <Socrates, wisdom> combine with instantiation in a way that <Socrates, foolishness> does not?

Individual qualities are intended to forge the missing link between universals and particulars. For them to play this role, it is essential that individual qualities not be conceived of as separable, either ontologically or conceptually, from the individuals that have them. If the wisdom of Socrates were thought of as potentially separate from Socrates himself, then we would require yet another item to account for the linkage between Socrates and his wisdom. The specter of the third man would haunt us still.

We may think of an individual quality as a part of the individual that has it, but it is a peculiarly inalienable part. We can imagine Socrates' nose continuing to exist even after Socrates had ceased to exist, but it is nonsensical to imagine Socrates' wisdom existing without Socrates. In this respect, individual qualities of particular individuals are like the surfaces of bodies. The surface of a body cannot exist without the body,[4] and the surface of a particular body cannot be the surface of any other body,[5] though it can resemble the surface of some other body. The body might have had a different surface – it might have been sanded down and painted – but the surface could not have had a different body. Likewise, the whiteness of this piece of paper could not exist without the piece of paper, and the whiteness of this paper cannot be identical to the whiteness of any other body, though some other body can exactly resemble this piece of paper in color. The piece of paper might have had a different whiteness – the wood pulp might have been bleached less thoroughly, for instance – but the whiteness of this piece of paper could not have belonged to anything else.

Our notion of individual quality is ultimately derived from Aristotle, who says

> Of things that are,...some are in a subject but not said of any subject (By 'in a subject' I mean what is in something, not as a part, and cannot exist separately from what it is in.) For example, the individual knowledge-of-grammar is in a subject, the soul, but is not said of any subject; and the individual white is in a subject, the body (for all color is in a body), but is not said of any subject.[6]

Recall Aristotle's view about geometrical objects.[7] Unlike Plato, for whom geometric objects were similar to forms, separate entities in a separate realm, Aristotle thought that the geometer studies the same entities that the physicist studies. Both study physical bodies, but while the physicist concerns himself with all aspects of the bodies he studies, the geometer is only interested in their size and shape, so he thinks about the bodies abstractly, mentally isolating those aspects of the bodies that he

cares about. Like the botanist, the geometer studies grapefruits, but he is only interested in grapefruits *qua* spherical and not in grapefruits *qua* sour or yellow or heavy. The geometer does not study separate abstract entities, he studies concrete entities regarded abstractly. In the language we are using, the sphericity of the grapefruit is one of its individual qualities. The concept of the individual sphericity of the grapefruit is obtained by abstraction, a mental operation by which one conceptually isolates those aspects of a thing one cares about.

So far, we have only talked about individual qualities, but there are also individual relations. There have to be, for we could never recognize an individual as tall just by attending to her individual qualities, but only by comparing her to others.

Among the individual relations are individual similarities. Individual similarities are useful in understanding generalization, the operation complementary to abstraction by which we group things together under a universal. Typically, we group things together that seem to us to be similar to one another in some respect. Similarity is not transitive – a may be similar to b and b to c without a being similar to c – so our grouping things under a general term cannot be simply a matter of forming individual – similarity classes. Most commonly, we classify things by picking a prototype of each class, and placing into the class those things that are sufficiently similar to the prototype and outside the class those things that are sufficiently different from the prototype, with some fuzziness around the border.[8]

There are exception to this general pattern, for sometimes we group together things that are not at all similar. Reid's example is the concept of a felony. An armed assault and a theft by computer are not objectively similar acts, yet we group them together under the general term "felony."[9] We do so because it is useful to do so. How we group things under general concepts is determined by considerations of utility. Most of the time, we group together things that are similar because most of the time it is useful to group together things that are similar. But we can do otherwise when it suits our purposes.[10]

As every reader of fiction knows, we can conceive of things that do not exist. Our theory of how we form general concepts tells us that we conceive of three kinds of things, particulars, individual concepts, and universals, but this does not, in itself, guarantee that these things exist. We now turn to the ontological question whether particulars, individual qualities, and universals exist.

Reid was no skeptic, and neither are we. Thus we want to affirm without reservation that particulars exist. Our opposition to skepticism also leads us to affirm that individual qualities exist, for to deny them would lead to a kind of Kantian idealism, where things in themselves lie forever outside the reach of our knowledge. If the whiteness of this piece of paper is not real, then it must be fictitious, and likewise all the qualities we attribute to a thing must be fictions, products of our minds. But if the qualities we attribute to things are all really just products of our own minds, then we can never really know how things are in themselves.

The question whether universals exist is more delicate. Reid himself was unequivocal. Universals do not exist. We conceive of universals – that is, according to Reid, we know the meanings of general terms – but when we conceive of universals, as when we conceive of centaurs, we are conceiving of something that does not exist. He thought this for at least two reasons. First were general considerations of ontological parsimony. Second was the observation that how we group things under general terms is determined by considerations of utility. But if our conceptions of universals corresponded to something objectively real, they would not depend on human purposes.

Notice that considerations of parsimony do not lead Reid to repudiate individual qualities. This is because an individual quality is not an additional item in one's ontology, separate from the particular that has it. To think of talk about individual qualities is to think or talk about particulars in a way that makes it easy to isolate certain aspects of the particulars for special attention. To deny the existence of individual qualities is not to advocate genuine ontological economy; it is merely to forbid a useful way of thinking and talking about particulars.

More cautious than Reid, we do not wish to deny the existence of universals. We merely want to say, with Reid, that we do not require the existence of universals in order to give an account of how we group things under general concepts; we require that there be a conception of universals, but not that there be anything answering to that conception. But we leave open the possibility that the existence of universals may be required for some other reason; for example, we may need to employ universals in giving the truth conditions for attributions of general terms.

Grouping things under general concepts is a mental activity we can perform well or badly. To some extent, this can be explicated counterfactually: the classification we made when we were careless, drunk, or observing by twilight is not the classification we would have

made if we had been careful, sober, and observing in broad daylight. We can introduce the notion of universals to represent classifications we would make if our mental apparatus worked flawlessly. What classifications are correct is not decided democratically. As Reid remarked, our conception of universals is a conception of how those most expert in some subject would classify things, for example, whether a lawyer or judge would classify a misdeed as a felony.[11]

Thus the universals describe the classifications that would be made by an ideal observer who employs the same classificatory principles that we employ or (in cases requiring technical expertise) that the experts employ, and who never falters in their application. The universal redness encompasses those things the ideal observer would judge to be red. We could use these universals to describe how an ideal observer would classify things, and we could even use them to describe how we classify things if we make proper allowances for our mistakes. But the universals would not be useful in *explaining* how we classify things. They would not even be useful in explaining how an ideal observer classifies things, since they would only tell us the results of the observer's classificatory judgments without telling up how the observer arrived at those judgments. Universals describe the classification our classificatory principles ideally aim at, but they do not explain how we carry out the classification. They give a teleological account of our classifications, rather than a casual one.

Suppose we want to explain how a certain computing device operates. For this purpose, it will not suffice to give a table that enumerates the output that an idealized error-free version of our machine would give for every input. Indeed even if our machine never makes any errors, so that the table describes its outputs precisely, the table would still not a provide a satisfactory explanation; for the table only lists the outputs, without telling us how these outputs are produced. The table might be useful descriptively, but it would have no explanatory value.

In describing or explaining how the mind works, universals play a role similar to that of the input-output table. We must distinguish the rules and procedures we use in making classificatory judgments from the classifications those rules and procedures would produce if faultlessly applied. Talk about universals is useful in describing, in an idealized manner, our classificatory judgments, but, if we want to explain those judgments, the universals will not be enough; we need to know the rules and procedures. But once we know the rules and procedures, we shall have no need for the universals, since the classifications are determined directly

by the rules and procedures, without bringing universals into the picture at all. So universals are neither sufficient to explain nor necessary to describe how we group things under general terms.

Universals are convenient for describing our classificatory system, but they are not essential for doing so. Whether these considerations of convenience are enough to override the countervailing concern for ontological parsimony is a question we shall not attempt to settle here.

Leaving the ontological status of universals thus unresolved, let us proceed to use the language of modal logic to describe some of the essential characteristics of individual qualities.

First of all, an individual quality is individual; at most one thing can have it. Our first attempt at spelling this out is as follows:

$\Box(\forall q)$(q is an individual quality $\rightarrow \Box(\forall x)(\forall y)$((x has q & y has q) $\rightarrow x = y$)).

But this does not quite work, since the yellowness of a grapefruit is the same as the yellowness of its rind. In its place, we have the weaker principle:

(Q1) $\Box(\forall q)$)q is an individual quality $\rightarrow \Box(\forall x)(\forall y)$((x has q & y has q) $\rightarrow$ x and y overlap mereologically)).[12]

Notice that this gives a necessary but not a sufficient condition for a quality to be an individual rather than a general quality. Being the tallest mountain is a quality that can be had by at most one thing, yet it is not an individual quality.

Every individual quality is a quality of some (existent) particular:

(Q2) $(\forall q)$(q is an individual quality $\rightarrow (\exists x)$x has q).[13]

This is because no nonexistent (e.g., fictional) character has existent individual qualities. A fictional horse might be just like Traveler in color, but a fictional horse could not have Traveler's color. Conversely, every particular has some individual qualities:

(Q3) $\Box(\forall x)$(x is a particular $\rightarrow (\exists q)$(q is an individual quality & x has q)).

An individual quality is ontologically dependent upon the particular that has it:

(Q4) $\Box(\forall x)(\forall q)$((q is an individual quality & x has q) $\rightarrow$ $\Box$(q exists $\rightarrow$ x exists (or, at least, part of x exists))).[14]

The parenthetical qualification "or, at least, a part of x exists" was added to account for the observation that, when we eat the grapefruit and throw away the rind, the grapefruit no longer exists but its individual yellowness does still exist, since the yellowness of the grapefruit is the same as the yellowness of its rind, and the rind still exists, its color unchanged.

The converse of (Q4) does not hold. The piece of paper might have existed without its whiteness existing, since the piece of paper might have been dyed a different color. But, necessarily, if the whiteness exists, the piece of paper has it:

(Q5) $\Box(\forall x)(\forall q)((q \text{ is an individual quality } \& \ x \text{ has } q) \rightarrow \Box(q \text{ exists} \rightarrow \text{some part of } x \text{ has } q))$.[15]

The consequent reads "some part of x has q" rather than "x has q" to take account of the following example, due to M. David: Let q be the yellowness of the grapefruit, and suppose that the outer layer of the rind is carefully removed. Then q still exists, for q is the yellowness of the outer rind, and the outer rind still exists, but the grapefruit no longer has q, for the grapefruit is no longer yellow.

Combining (Q1) and (Q5), we get:

(Q6) $\Box(\forall x)(\forall q)((q \text{ is an individual quality y } \& \ x \text{ has } q) \rightarrow \Box(\forall y)((q \text{ exists } \& \ y \text{ has } q) \rightarrow x \text{ and } y \text{ overlap merelogically}))$.[16]

D. Schulthess has proposed an interesting counterexample to this last principle. Let us imagine that we see a large white dog in a field, but we mistake it for a sheep. In supposing that there is a sheep in the field, we believe something false, but we do not believe anything crazy. Is it not possible for there to have been a sheep in the field that has exactly the same color that the dog in fact has?

We respond to this example by invoking Kripke's distinction between epistemic and metaphysical necessity.[17] It is epistemically possible for the dog to have been a sheep, in which case the sheep would indeed have had the color the dog in fact has. But it is not metaphysically possible for the dog to have been a sheep nor for the whiteness of the dog to have been the whiteness of a sheep. All that is metaphysically possible is for there to have been a sheep in the field that exactly resembles the dog in color. All our "$\Box$" 's are to be understood as attributions of metaphysical necessity.

These modal theses describe the metaphysical status of individual qualities, which, though distinct from the particulars in which they reside,

are not separable from the particulars in which they reside. There is one important feature of the metaphysical status of individual qualities that we have not yet discussed, namely, that what makes an individual quality distinct from the particular in which it resides is the active intercession of the human mind.

What is it that distinguishes an individual quality of a particular from the other individual qualities of the same particular and from the particular itself? Nature distinguishes the rind of the grapefruit from the flesh and the seeds, but what distinguishes the yellowness of the grapefruit from its roundness is not nature but mind. We attend to the yellowness of the grapefruit separately from its other aspects, and so we distinguish the yellowness of the grapefruit from the roundness of the grapefruit and from the grapefruit itself. Consider the taste of the grapefruit. The distinctive taste of the grapefruit is the effect of an exceedingly complex combination of its individual biochemical properties. The taste of the grapefruit is one of its individual qualities, yet there is nothing in the internal structure of the grapefruit that distinguishes this particular combination of biochemical properties from myriad other biochemical features of the organism. What distinguishes this particular combination of biochemical properties is us; this is the combination of properites to which our sensory apparatus responds. The taste of the grapefruit is a consequence not only of what the grapefruit is like, but of what we are like.[18]

The taste, color, and shape of the grapefruit are all aspects of an object that exists independently of us and our thinking. But what distinguishes these aspects from each other and from the grapefruit itself is our thinking about them. What effects the transition from grapefruit to grapefruit *qua* sphere is the intensional and intentional activity of the human mind.

Individual qualities are aspects of particulars, and most particulars exist independently of human minds and thought. Yet what individuates individual qualities, what distinguishes them from each other and from the particulars in which they inhere is the action of our minds. Having this dual nature, being in part mind-dependent and in part mind-independent, individual qualities occupy a position intermediate between particulars, which are mind-independent, and universals, which are, on a Reidian account, constructions of the human mind. It is this intermediate position that enables individual qualities to play a central role in the theory of predication.

NOTES

[1] Thomas Reid, *The Works of Thomas Reid*, D.D., Eighth Edition, Sir William Hamilton, ed., (Edinburgh: James Thin, 1895), pp. 360-412. For an abbreviated edition of Reid's work treating the subject of this paper, see Thomas Reid's *Inquiry and Essays*, Ronald E. Beanblossom and Keith Lehrer eds., (Indianapolis, Hackett, 1983), pp. 1-296, *passim*. References to Reid are given to the edition of Reid's work that is most frequently found in libraries, though it is not the most faithful edition. Research on Reid by Keith Lehrer was supported by a grant from the National Science Foundation and a John Simon Guggenheim Memorial Foundation Fellowship. We wish to thank Marian David and Scott Sturgeon for their comments on the penultimate version. Although our proposal is based on the work of Reid, many subsequent authors have advanced the notion of individual qualities, most notably Peter Simons, Barry Smith and Kevin Mulligan in their paper, "Truth-Makers", *Philosophy & Phenomenological Research*, 1984, 287-321 and their contributions to *Parts & Moments. Studies in Logic & Formal Ontology,* Munich: Philosophia, 1982.

[2] Ibid., pp. 394-5.

[3] Ibid., p. 395.

[4] Of course, a hollow body still has a surface, and the body we are looking at might have had the same surface if it had been hollowed out.

[5] Except in the sense that the surface of a grapefruit is the same as the exterior surface of the rind of the grapefruit.

[6] *Categories* 1a20-29. J.L. Ackrill's translation (Oxford: Clarendon Aristotle Series, 1963).

[7] See especially *Physics* 193b31-194a13.

[8] This account of how we arrive at general concepts is derived from the work of Eleanor Rosch.

[9] Reid, p. 364.

[10] Ibid., pp. 394-98, 405-10.

[11] Ibid., p. 395.

[12] There is an analogous principle governing individual relations, namely:

(R1) $\Box(\forall r)$(r is an n-ary individual quality $\rightarrow$ $\Box(\forall x_1)\ldots(\forall x_n)(\forall y_1)\ldots(\forall y_n)((\langle x_1,\ldots,x_n\rangle$ has r & $\langle y_1,\ldots y_n\rangle$ has r) $\rightarrow$ (x_1 and y_1 overlap mereologically & ... & x_n and y_n overlap mereologically))).

[13] The relational analogue is:

(R2) $\Box(\forall r)$(r is an n-ary individual relation $\rightarrow$ $(\exists x_1)\ldots(\exists x_n)\langle x_1,\ldots,x_n\rangle$ has r).

[14] The relational analogue is:

(R4) $\Box(\forall x_1)\ldots(\forall x_n)(\forall r)$[(r is an n-ary individual relation & $\langle x_1,\ldots,x_n\rangle$ has r) $\rightarrow$ $\Box$(r exists $\rightarrow$ (at least part of x_1 exists & ... & at least part of x_n exists))).

[15] The relational analogue is:

(R5) $\Box(\forall x_1)\ldots(\forall x_n)(\forall r)$ ((r is an n-ary relation & $\langle x_1,\ldots,x_n\rangle$ has r) $\rightarrow$ $\Box$(r exists $\rightarrow$ $(\exists y_1)\ldots(\exists y_n)$($y_1$ is part of x_1 &...& y_n is part of x_n & $\langle y_1,\ldots,y_n\rangle$ has r)).

[16] The relational analogue is:

(R6) $\Box(\forall x_1)\ldots(\forall x_n)(\forall r)[(r \text{ is an n-ary individual relation } \& <x_1,\ldots,x_n> \text{ has } r) \rightarrow \Box(\forall y_1)\ldots(\forall y_n)((r \text{ exists } \& <y_1,\ldots,y_n> \text{ has } r) \rightarrow (x_1 \text{ and } y_1 \text{ overlap mereologically } \&\ldots\& \ x_n \text{ and } y_n \text{ overlap mereologically}))]$.

[17] See pp. 319ff of "Naming and Necessity" in Donald Davidson and Gilbert Harman, eds., *Semantics of Natural Language* (Dordrecht, Holland: D. Reidel, 1972), pp. 253-355.

[18] According to Quine and Goodman ("Steps Toward A Constructive Nominalism," *Journal of Symbolic Logic* 12 (1947): 105-22), there is no difference in ontological status between Cleopatra and the being with Cleopatra's head and Caesar's torso. The reason why we would ordinarily think of Cleopatra as a single individual and Cleopatra's-head-with-Caesar's-torso as an artificial construct is not any objective difference in their ontological status, merely that we find it useful to think of the former but not the latter as a single unit. While Quine and Goodman's position is counterintuitive with respect to particulars (this is not to say it is mistaken), the corresponding position with respect to individual qualities – what combinations of features of a thing we regard as constituting individual qualities is determined by our beliefs, desires, and perceptions – is highly plausible.

BARRY SMITH

CHARACTERISTICA UNIVERSALIS

1. PREAMBLE

Our task will be to construct portions of a directly depicting language which will enable us to represent the most general structures of reality. We shall draw not on standard logical treatments of the contents of epistemic states as these are customarily conceived in terms of propositions. Rather, we shall turn to a no less venerable but nowadays somewhat neglected tradition of formal ontology, in which not sentences or propositions, but maps, diagrams or pictures, shall serve as the constituents of our mirror of reality.

The construction of a directly depicting language we conceive as being in a certain sense an experimental matter. One can, as Peirce remarked, 'make exact experiments upon uniform diagrams', and in formal ontology as we conceive it operations upon diagrams will 'take the place of the experiments upon real things that one performs in chemical and physical research.'[1] The chemist is of course grappling in his experiments with the very bits of reality whose properties he is concerned to establish. Here, in contrast, we shall be experimenting at one remove, with *diagrams* of reality. The contrast is not so great as it might seem, however. For it is not, in fact, the particular samples in which the chemist is interested; rather (in Scotist vein):

> the object of the chemist's research, that upon which he experiments, and to which the question he puts to Nature relates, is the Molecular Structure, which in all his samples has as complete an identity as it is in the nature of Molecular Structure ever to possess.

And now, as Peirce once more rightly insists, it is not otherwise with experiments made upon diagrams. The latter are 'questions put to the

Nature of the relations concerned' – precisely in virtue of the fact that we are here experimenting with diagrams which are to enjoy the property that the forms of relations exemplified in reality will be the very same as the forms of relations in the diagrams themselves.

A similar idea is of course present also in Wittgenstein. As the *Tractatus* has it: 'What constitutes a picture is that its elements are related to one another in a determinate way.' (2.41) Indeed: 'There must be something identical in a picture and what it depicts, to enable the one to be a picture of the other at all.' (2.16) Wittgenstein's 'pictorial form', then, is Peirce's 'form of a relation', and our task here will be one of taking further the idea of a universal characteristic which both philosophers shared.

2. FROM LEIBNIZ TO FREGE

The project of such a characteristic had of course been envisaged by Leibniz, and the idea is present already in Descartes and Jungius and in others before them.[2] In each case we have the idea that it is possible to isolate a relatively small number of basic units and of structure-building principles governing the combination of such units in a way which will allow, by more or less mechanical means, the felicitous representation of all the concepts, truths, thoughts, or structures, pertaining to a given sphere. The crucial idea here is that of compositionality. Already Jungius saw the mathematical method for the study of nature as valid precisely because

> nature does not act the way the Chinese write, but like other peoples, i.e., with an alphabet ... through combinations, complications, and replications of a few hypotheses, laws, or principles[3]

In the work of Descartes, Jungius and Leibniz, however, two distinct ideas are run together: the idea of the characteristic as a perspicuous representation of relations among concepts, and the idea of the characteristic as a mirror of reality. Only in the brief realist interregnum around the turn of the present century, and especially in the work of Brentano and his followers, in the *Tractatus*, and to some extent also in Peirce, did the second of these two ideas, the idea of an ontological characteristic, come into its own.

Brentano himself applies this idea to the specific field of psychology. Brentanian descriptive psychology is accordingly an example of a *characteristica specialis*, a directly depicting language restricted to some

specific sphere. That he is seeking an ontological characteristic in our sense is nonetheless clear. He describes the discipline of descriptive psychology as having been conceived as an instrument that would

> display all the ultimate psychic components, from whose combination one with another the totality of psychic phenomena would result, just as the totality of words is yielded by the letters of the alphabet.[4]

The project of a *characteristica universalis*, of a directly depicting language that would apply to all spheres of reality without restriction, is sketched by Husserl in the *Logical Investigations* – a work which can in many respects be conceived as an extrapolation of Brentanian ideas in the direction of a completely general formal ontology. Husserl provides however no more than philosophical preliminaries to the working out of the formal details of a characteristic. The first sophisticated moves in this formal direction after Peirce are provided, rather, in the work of Frege, who had the key idea underlying the project of a directly depicting language that the syntactic relations of expressions in such a language should mirror exactly corresponding ontological relations among the entities depicted. Moreover, Frege first showed the possibility of a new sort of syntactic precision and determinacy in this respect – specifying for the first time a language with a definite notion of well-formedness. The expressions of Frege's system are, familiarly, built up via the nexus of function and argument: one expression is said to 'saturate' another (incomplete) expression, as when we move from 'the square root of' to 'the square root of 2'. (The chemical connotations of Frege's terminology of 'saturation' here are not accidental.)

In some of his earlier writings, now, Frege suggests that the complex expressions of his system should map corresponding structures among saturated and unsaturated entities in the world, in such a way that, when a function is saturated by an argument, then what results is a whole of which the function is a part. (Wittgenstein, too, seems to presuppose a similar idea e.g. at 5.47, where he suggests that a thing's being composite involves in every case function and argument.) Certainly it may be argued that mereological relations are preserved in this way when functional expressions are saturated by other expressions of an appropriate type. The expression 'father of' is in a certain sense part of the more complex expression 'father of John'. A moment's reflection makes it clear, however, that such mereological relations on the syntactic level cannot, on the Fregean approach, mirror corresponding relations among the

objects in the world – John himself is not a part of his father, any more than Denmark, say, is a part of Copenhagen. Thus as Dummett points out, Frege quickly saw that it had been wrong for him ever to have suggested that a parallelism of the given sort can be maintained.[5] The most that can be affirmed is that the unsaturatedness of a functional expression mirrors a corresponding unsaturatedness in the function for which it stands.

Frege does hold on to the idea of a characteristic language in certain respects, however. Above all, the absence of semantics in Frege's logic puts him firmly within the tradition of characteristic ontology as here conceived. A truly adequate directly depicting language does not need a semantics, for the relation of language and world is here able to take care of itself. Frege's clear departure from the characteristic tradition is seen, however, in his insistence that all expressions of his system, whether saturated or unsaturated, are referential – so that he spurns the admission of empty complex names (names which would refer to nothing at all because the corresponding objects do not stand to each other in the way they are represented as standing). Frege's work is in this respect similar to that of Meinong, though the similarity is masked by Frege's tendency to identify the referents of names that would intuitively count as referring to what is non-existent – names like 'the grandmother of Denmark' – with some arbitrarily chosen dummy such as the False or the number 0. For other reasons, too, the names of Frege's language (and above all most names of the True and the False), do not mirror the structures of what they denote or represent. At best we might say that they mirror the structure of a certain method for fixing or identifying their respective denotata.

3. DIRECTLY DEPICTING DIAGRAMS VS. EXISTENTIAL GRAPHS

The failure of Frege's system as the basis of a characteristic flows, we might say, from the fact that he confused grammar and ontology. This is because Frege allowed intuitions deriving from investigations of grammatical well-formedness to determine his account of the structures of the objects depicted by the diagrams of his system. (Meinong allowed his ontology to be determined in a similar way by psychological considerations.) In a properly constructed characteristic language, in contrast, the relation of determination will flow in the opposite direction: the structures

of diagrams will as far as possible be dictated by the structures in reality they are designed to represent. In particular, the directly depicting language to be constructed in what follows will hold as doggedly as possible to the idea that the part-whole relations among its constituent expressions must map exactly corresponding relations among the pictured objects. Thus we are interested, as Peirce, again, expressed it:

> in constructing an icon or diagram the relations of whose parts shall present a complete analogy with those of the parts of the object of reasoning, of experimenting upon this image in the imagination, and of observing the result so as to discover unnoticed and hidden relations among the parts. (3.363)

Peirce was however interested above all in constructing diagrammatic languages for purposes of logic (though 'logic', for him, as for Frege, overlapped to some extent with what is here called 'formal ontology'). Peirce's existential graphs were designed to 'render literally visible before one's very eye the operation of thinking *in actu*.' (4.6)[6] Here, in contrast, we are interested in the idea of constructing a diagrammatic language for purposes strictly ontological. Like Peirce we shall employ a modification of Euler's diagrams to this effect.[7] Peirce saw it as a defect of such diagrams as originally conceived that they relate essentially to dichotomous relations among what is general, i.e. to relations among the ontological correlates of common names of the sort that are at issue for example in classical Aristotelian syllogistic. Thus the inclusion of one ovoid α inside another β might indicate that the concept α is subordinate to the concept β as *red* is subordinate to *colour*. Such diagrams operate entirely in the sphere of what is general. Hence they cannot be used to affirm the existence of any individual satisfying a given description. This may suffice for the purposes of the logician. The formal ontologist, however, is concerned with the depiction of what actually exists, and since he is concerned not merely with what is general but also with what is particular, his diagrammatic language must embrace also what we might call indexical or 'proper' diagrams, which will function in much the way that proper names function in ordinary language.

Our diagrams will in fact almost all of them incorporate proper names in the ordinary sense as constituent parts. It is from these that they will inherit their primary relation (be anchored) to reality. It will simply be assumed in what follows that such ordinary proper names refer; we shall not ask in virtue of what they refer, though we note that this would be a proper and pressing question for another occasion. We shall say that a

diagram is true if that which it sets out to depict exists in reality. As 'elementary diagrams' we shall allow proper names standing alone; and it will turn out that, because empty proper names are not admitted, every elementary diagram of this sort is true. A complex diagram is true if and only if the structural relations between the names (and other bits and pieces in the diagram) map structural relations among the corresponding objects in the world. Otherwise it is false. All of this should, of course, be perfectly familiar.[8] Our diagrams will furthermore be synchronous only. Thus by 'the world' or 'reality' we shall mean reality as it is now, at the present time. The treatment of time and change from the standpoint of a universal characteristic we postpone for the future.

4. SOME CONDITIONS ON A DIRECTLY DEPICTING LANGUAGE

It is above all in the ideal language philosophy defended by Wittgenstein in the *Tractatus*, by Gustav Bergmann and his followers, and by Cocchiarella and others in the present day, that the project of a *characteristica universalis* has received its most sophisticated modern expression. Even these thinkers, however, too often run together ontological and logical concerns. Most importantly, they identify logical and ontological simplicity and complexity, and thereby also they assume that logical atomism implies one or other form of ontological atomism. Hence the project of a directly depicting 'ideal' language has come to be associated with atomistic doctrines of one or other sort as concerns the objects which such a language might depict. Consider, for example, Russell's atomism of thises and thats, Bergmann's doctrine of bare particulars, or Grossmann's modification of this doctrine in his *The Categorial Structure of the World*. Here, in contrast, the idea of a directly depicting language built up out of proper names in something like the Tractarian sense will be combined with a view of such proper names as referring to common-or-garden substances like Chisholm or this desk, and to the various different sorts of parts and moments thereof. Proper diagrams, then, to the extent that they are true (i.e. are such as to depict existent structures in reality) will depict facts involving ordinary objects of these and similar sorts. Whether, or to what extent, the language can be extended to cope also e.g. with microphysical structures is a question we here leave open.

A directly depicting language seeks high representational adequacy, even at the expense of low expressive adequacy. Thus our language will

have no facility to represent propositional attitudes, it will have no probabilistic machinery, and it will tolerate no vagueness. Work in formal philosophy of the last 70 years or so has concentrated overwhelmingly on problems of the representation of *knowledge*, which is to say on the various ways in which partial and sometimes incorrect information is acquired, stored and processed. Here, in contrast, we are interested not in the peculiarities of *episteme* and *doxa*, but purely in the idea of a direct depiction of reality taken for its own sake, which is to say independently of any concern as to how such direct depiction may play a role in our common concerns with reality as knowing subjects.

Valuable groundwork in this respect is to be found in Gustav Bergmann's "Notes on Ontology" of 1981. Here Bergmann formulates four conditions on what he calls an 'improved language', conditions we here reproduce in somewhat simplified form:

(F1) With four exceptions, every primitive mark of the improved language stands for an existent.
(F2) Every well-formed 'string' of the improved language stands for a determinate, i.e., for either a thing or a complex or a class.
(F3) Every primitive standing for a thing and every well-formed string of the improved language stands for one and only one existent; and, with a single exception (relating to classes), conversely.
(F4) (a) There are in the improved language no complex characters such as, say, green-and-square. In the improved language, therefore, there are, by (F1), (F2), no derived predicates, whether introduced by an abstraction operator or (as one says, metalinguistically) by definitions.
(b) There are by (F3) in the improved language none of the expressions Russell called incomplete symbols. To epitomize both (a) and (b) in a way that throws some light on the reasons for them, the improved language contains no abbreviations.
(c) In the improved language there are no variables. Any notation using them prevents one from arriving at an ontologically adequate assay of quantification.

Ad (F1): The problem of depicting putative intentional facts such as are expressed, e.g., by sentences like 'John is thinking of a unicorn' is one central *experimentum crucis* for a directly depicting language. Accordingly the most important exception of the four mentioned by Bergmann in the first of his four conditions is the primitive mark '**M**' (roughly: 'is thinking of'), which for Bergmann is an ingredient of the expression of the connection between a thought and its intention. **M**, in Bergmann's view, 'stands literally for nothing' (p. 140). The diagrammatic language to be presented below, in contrast, satisfies Bergmann's first condition *with no exceptions at all*. On the other hand, this language may

be weaker than that projected by Bergmann, in that it is not capable of expressing facts of the sort Bergmann expresses using '**M**' (if indeed there are such facts).

Ad (F2): Bergmann, as he himself tells us, uses the word 'string' merely provisionally, since he had come to the view that no linear notation can accommodate an ontologically adequate assay even of linear order. As it happens he, too, suggests an Euler-type notation for his improved language, in which for example the exemplification of what he calls a 'character' (for example green) in an object is represented (roughly) by inscribing the name of the object in a circle and the name of the 'character' in the space between this circle and a larger concentric circle drawn around it.[9]

Ad (F3): We, too, shall accept that every (true) diagram of our language will stand for one and only one existent. On the other hand, there are for us many cases where a single existent admits of being truly diagrammed in a multiplicity of distinct ways. Perhaps we can say that Bergmann was able to take seriously the goal of one diagram per entity because entities, for him, are ultimately very simple indeed. Here, on the other hand, we are interested in establishing a diagrammatic language that is able to provide different snapshots of one and the same segment of reality as it were on different levels or with different finenesses of grain, so that different sorts of detail are brought into focus. (Compare the way in which incompatibilities arise among different Fregean analyses of linguistic objects into saturated and non-saturated parts.) The ontologist then has a means of representing certain non-depictable facts positively, by allowing meta-diagrammatic assertions, not officially a part of the diagrammatic language as such. That is, he can express such facts by means of assertions about the diagrams themselves. Thus for example he can assert that two entities are identical by means of a statement like $\Theta = \Sigma$, where Θ and Σ are diagrams of the entities in question. He might consider also for example a substitution law to the effect that, if '$\Theta = \Sigma$' is true, and if '$\Gamma(\Sigma/\Theta)$' is like 'Γ' except that 'Θ' has been replaced by 'Σ' in one or more places, then from Γ one may infer $\Gamma(\Sigma/\Theta)$.

Ad (F4)(a): We, too, in what follows, shall reject all derived terms for what is general, whether introduced by abstraction or by definition.

Ad (F4)(b) We shall also allow no Russellian incomplete symbols. However, since we allow different (true) diagrams of one and the same existent, and since some of these diagrams will turn out to be simpler than others, our language does in this sense incorporate a facility for something

like abbreviations. Abbreviations, for us, are possible because of our caveat to (F3); but still, we accept nothing like definite descriptions.

Ad (F4)(c) We agree with Bergmann also in the rejection of variables. Reality is determinate, so that any notation capable of directly depicting this reality must also be determinate.[10] We shall, though, allow metavariables which will enable us to make assertions about our diagrams themselves.

5. THE OIL-PAINTING PRINCIPLE

Bergmann's four conditions are interesting and important, but he and his followers have been blinded by the successes of Fregean logic to the extent that their conceptions of an ideal or improved language still reveal elements of compromise as between the logical and epistemological purposes of knowledge-representation and the ontological purposes of object-representation appropriate to a directly depicting language as here conceived. Above all, Bergmann and his followers, in neglecting the mereological constraints at the heart of the idea of a directly depicting language and in seeking to hold too closely to the paradigm of Fregean logic, did not see that the principal mark of a characteristic language is that it satisfies:

> **The Oil-Painting Principle (also called 'Degen's Law')**: If diagram Θ is a well-formed part of diagram Γ, then if Γ is true, so also is Θ.[11]

This, too, is a metadiagrammatical principle affirming that one diagram can be inferred from another. A formulation of the intuition underlying the principle – an intuition shamefaced in its naivety – might be: every part of a representative work of art is also representative. Clearly, in a language which satisfies a principle of this sort, there can be no disjunctive, negative or hypothetical diagrams.[12] Already at this level, therefore, the purposes of logic and the purposes of direct depiction are alien to each other, a fact that was perhaps recognized by Wittgenstein in the *Tractatus*, e.g. in his drawing of the distinctions between *Elementarsatz* and *Satz* and between *Sachverhalt* and *Tatsache*. The relation of making true, here, holds exclusively between the *Elementarsatz* and the *Sachverhalt*: the truth of *Sätze* is an entirely derivative matter. If, now, we define logical atomism as a view to the effect that only that can serve to make a sentence or proposition true which can be depicted in a directly depicting language,

then it would follow that there are, for the logical atomist – and as Wittgenstein correctly saw – no disjunctive, negative or hypothetical states of affairs.

Can we accept a converse of the oil-painting principle, to the effect that, if Θ and Γ are any true diagrams, then there is a true diagram containing Θ and Γ as parts? This would ensure that diagrams constitute at least an upper semi-lattice on the obvious mereological ordering, with unit something like a complete diagram of the world. Or can we accept an even stronger principle, amounting to the affirmation of a unity of reality of an almost Spinozistic sort, which would assert that there is some true *connected* diagram containing any Θ and Γ as parts? Problems will arise for any such law in virtue of the fact that we could not, in general, read off from the antecedently given diagrams what the relevant containing diagram (least upper bound) would be. Thus we should have no effective way of setting such a law to work. This, though, may be merely an epistemological matter, of no ultimate significance for ontology.

On the other hand it seems clear, at least on our intuitive understanding of the ways in which the bits of worldly furniture are related, that we can on no account embrace a general intersection law which would allow us to infer, from any given pair of diagrams, to their intersection or meet. For consider some pair of diagrams depicting disjuncta (in the sense of entities that have no parts in common). The intersection of disjuncta is, by hypothesis, empty. Yet it seems clear that a directly depicting language can on no account allow what might be called an empty or null diagram, since such a diagram would, by definition, depict nothing at all. Thus the system of diagrams cannot constitute a lower semi-lattice. Mereology, too, allows nothing like a null element, and thus the mereological intuitions at the heart of the idea of direct depiction make themselves once more clearly felt.

6. PRIMITIVES AND DEFINITIONS[13]

The goal of a directly depicting language is that of constructing a system of diagrams that will allow the direct and adequate depiction of a maximum number of (kinds of) ontological facts. As we have seen, some sorts of facts – for example all facts involving negation or disjunction – cannot be depicted directly. Identity, too, is a relation of this sort. Bergmann gives arguments for supposing that diversity is a positive fact,[14] and certainly some sorts of diversity (for example mereological discreteness) can be

directly depicted. Facts of identity would then be negative facts, which would make it understandable that they cannot be directly depicted.

How a finally acceptable directly depicting language will look, will clearly depend on what the world is like. Hence the route one takes towards the construction of such a language will reflect one's initial ontological intuitions, which might for example be of a Fregean, Machian, Tractarian, Chisholmian, or Bergmannian sort. In any case, it will be necessary to start with some one or other candidate world-picture and to experiment with the construction of a directly depicting language appropriate thereto. Perhaps standard predicate logic is the best one can achieve in the direction of a directly depicting language for a universe enjoying a certain sort of set-theoretic structure. The fact that standard predicate logic fails almost every test of direct depiction might then serve as a reason for abandoning set theory as a basis for ontology. In general, however, one's candidate world-picture will not, in this way, need to be abandoned entirely. Rather, it will have to be tentatively adjusted in light of problems encountered when the attempt is made to represent it in a diagrammatic way. Such adjustments may then lead in turn to new principles of direct depiction, and these may reflect back in the form of adjustments to the underlying ontology, in a cycle which may be repeated for as long as it takes to reach a match of language and ontology of the appropriate sort.

Before turning to the directly depicting language itself, then, we need to make clear the intitial repertoire of ontological categories and relations between categories to which the various bits and pieces of its machinery shall correspond, in order that the cycle of diagrammatic experimentation can begin. Here we shall follow Aristotle and some scholastics in taking as our initial focus the notion of an *individual substance* or continuant, examples of which would be such ordinary objects as human beings, oxen, logs of wood. Associated herewith is the Aristotelian category of individual accident (individual qualities, actions and passions, Rupert's present knowledge of Greek, an electric charge, a bruise, a blush). Accidents are said to 'inhere' in their substances, a notion which will be defined more precisely in what follows in terms of the concept of specific dependence. Substances and accidents together will constitute what we shall call the 'atoms' of our system of ontology.

'Individual', here, is a primitive term, and its precise meaning will become clear only in the course of what follows. As we understand the term, however, it is ruled out that there might be abstract individuals: to be individual is to be fully a part of the constantly changing order of space and

time. We do not, however, wish to presuppose that everything is individual: good candidate examples of non-individuals might be universals like *red*, entities capable of being exemplified by or realized in a multiplicity of individuals. Hence we shall recognize stocks of both individual and non-individual names, leaving open the possibility that the latter may turn out to be empty. Both individuals and non-individuals may be simple or complex. We adopt hereby a convention to the effect that a whole is individual if any part is individual (so that all the parts of a non-individual are themselves non-individual). Moreover, we shall assume that non-individuals are such as to exist only *in re*, which is to say that they exist only insofar as they are exemplified by or realized in individuals. This implies the following:

> **Weak Law of Immanent Realism**: If there is anything, then there is something individual of which it is a part.

We shall also assume, for present purposes, that no individual is such as to exist necessarily. From which it will follow that non-individuals, too, enjoy a merely contingent existence (they exist only for as long as, and to the extent that, there are individuals in which they are realized or exemplified.)

The notions now defined (in terms of the primitives 'individual', 'part',[15] and 'is necessarily such that') will be the formal or categorial notions by which the further construction of our diagrammatic language will be motivated and in terms of which it will be described. x, y, etc., are metavariables standing in for proper names of individuals and non-individuals. We define first of all:

> x is disjoint from y =df. x and y have no parts in common.

> x is discrete from y =df. x and y are individuals which have no individual parts in common.

Jules and Jim are discrete from each other in this sense. If, however, they contain as parts in common universals such as *human* or *animate*, then they are not disjoint.

To capture the notion of inherence, the relation holding between an accident and that which it is an accident of, we now put:

> x is specifically dependent on y =df. (1) x is an individual, and (2) x and y are such as to have no individual parts in common, and (3) x is necessarily such that it cannot exist unless y exists.

A headache, for example, is specifically dependent on me (as also on my head). This is a case of one-sided specific dependence (for it is clear that I am not specificially dependent on my headache). As we shall see, however, there are also cases where entities stand to each other in relations of mutual or reciprocal dependence.

We can now define:

> z is the mereological sum of x and y = df. (1) x is part of z, and (2) y is part of z, and (3) no part of z is disjoint from both x and y.

> x is a one-sidedly separable part of y = df. (1) x is a proper part of y, and (2) some part of y discrete from x is specifically dependent on x, and (3) x is not specifically dependent on any part of y discrete from x.

x is for example a thinker and y is the mereological sum of x together with some one of x's thoughts.

> x and y are mutually separable parts of z = df. (1) z is the mereological sum of x and y, and (2) x and y are discrete from each other, and (3) x is not necessarily such that any individual part of y exists and (4) y is not necessarily such that any individual part of x exists.

z is, for example, a pair of stones, and x and y the stones themselves. It can be seen to follow from the definition that only individuals are candidates for being either one-sidedly or mutually separable parts.

> x is atomic = df. (1) x is an individual, and (2) no part of x has either one-sidedly or mutually separable parts.

If x is atomic, we can infer that all proper parts of x stand in mutual dependence relations to other proper parts of x. The concept of an atomic entity does not yet capture the concept of substance however. For what is atomic need not, according to this definition, be independent or self-sustaining. Thus there might be dependent entities, for example in the realm of qualities, which are atomic by our definition. Accordingly we define:

> x is substantial = df. (1) x is atomic and (2) x is not specifically dependent on any other entity.

We are still not home, however, for our definition of 'substantial' is satisfied by quantitative parts of substances, such as Darius's (undetached) arm. The latter is atomic, because any proper part must share a boundary with some adjacent part and this would imply, as we shall see, a (mediate) dependence on the latter. But the undetached arm (or a growing tree) is clearly not a substance, either; for it, too, shares a boundary with some other entity, and is therefore once more dependent on something other than itself. What we loosely refer to as Darius's arm becomes a substance only on becoming detached, when it acquires a boundary of its own. Much of this will become clearer when once the relevant portions of the directly depicting language have been set down. For the moment we simply note that in order to arrive at a definition of substance we shall need to take account of the notion of boundary. To this end we introduce a new sort of dependence:

> x is boundary dependent on y =df (1) x is a proper part of y, and (2) x is necessarily such that either y exists or there exists some part of y properly including x, and (3) each individual part of x satisfies (2).

Thus for example the boundary of a billiard ball is a part of and is boundary dependent on the ball itself. We can now define:

> x is self-boundingly substantial = df. (1) x is substantial, and (2) there is no y that is boundary dependent on x and on some object discrete from x.

That Darius's undetached arm (or, again, a tree that is growing out of the ground) do not satisfy this definition will allow us to define substances as minimal self-boundingly substantial objects:

> x is a substance =df. (1) x is self-boundingly substantial and (2) no proper part of x is self-boundingly substantial.

Parallel to the distinction between 'substantial' and 'substance', we now have a distinction between 'accidental' and 'accident':

> x is an accidental of y =df. (1) x is atomic and (2) x is not substantial and (3) x is specifically dependent on y.

This definition has the useful property that it allows us to recognize that there are accident-like entities which relate to undetached or quantitative parts of substances as accidents proper relate to substances themselves.

(Consider the individual redness of the snake that is half red and half white.)

In what follows our attentions will be directed principally to accidents of substances. Hence we shall define:

x is an accident =df. x is an accidental of a substance.

We can now define:

x is an atom =df. x is either a substance or an accident.

The atoms are, as it were, the most privileged, most natural, most self-contained and well-rounded examples among the domain of atomic entities in general.

What is substantial is always part of some substance, and what is accidental is always part of some accident. From this it follows that to recognize the categories of substantials and accidentals is to add nothing new to the totality of what exists. It reflects cuts skew to those which pick out substances and accidents, and the latter, we suggest, reflect the most natural joints in reality. In what follows, accordingly, we shall ignore what is substantial as such, and see individual reality as being divided into substances and accidents alone. The world, then, is the totality of atoms, and the relation of specific dependence is the bond which holds these atoms together in *molecules* of different sorts.

To capture the latter notion we first of all define:

x is closed under specific dependence =df. no part of x is specifically dependent on any entity discrete from x.

(Everything substantial is closed under specific dependence in this sense.)

We can now say:

x is a molecule =df. (1) x is closed under specific dependence, (2) x contains a substance as a proper part, and (3) any pair of discrete parts of x are connected, directly or indirectly, by relations of specific dependence.

If, as seems reasonable, we exclude the possibility of what we might call lazy atoms, which is to say atoms which do not enter with other atoms into molecules of any sort,[16] then it would follow that the world is not only the totality of atoms; it is also the totality of molecules. Again, no contradiction arises here, since the two given assays of the totality of what exists reflect cuts at different levels. Unlike Wittgenstein (and Aristotle, and

Bergmann) we are not disturbed by the possibility of ontological inventories which reflect different sorts of modes or thicknesses of division in this way. Already every extended thing can be seen as being sliced along an infinity of different internal boundaries. This possibility will, however, imply that the idea of a single universal diagram is an idea that must be treated with caution.

Substances, as we have seen, may have substantials as proper parts. Accidents, correspondingly, may have accidentals as proper parts. Both substances and accidents may also, however, have certain essential parts, parts whose destruction leads to the destruction of the whole. Jim's individual humanity may be counted as an essential part of the substance Jim in this sense. Hue, saturation and brightness may similarly be taken as essential parts of that accident which is Jim's whiteness.

To capture this notion of essential part we shall introduce the terminology of 'sub-atoms'.

> x is a sub-atom =df. (1) x is a proper part of an atom and (2) x is not substantial or accidental.

Further examples of such essential parts are: the individual shape or mass of a body, the individual pitch, timbre or loudness of a tone. (If we entirely deprive a tone of its loudness we thereby destroy the tone itself.) As in the case of the hue, saturation and brightnesss of a colour, so also here, we are dealing with a case of three-fold mutual dependence. Indeed it turns out that many varieties of sub-atoms are entities which cannot as a matter of necessity exist except in consort with other sub-atoms of specific sorts.

Accidents may be relational or non-relational. They are relational if they depend upon a plurality of substances. Non-relational accidents are attached, as it were, to a single carrier, and thereby form, with their respective substances, the simplest kind of molecular whole. Relational accidents join a plurality of carrier-atoms together into more complex molecular wholes. Two or more individuals (be they atoms, sub-atoms or molecules) not joined together either by bonds or by further relational accidents constitute what we might call an *aggregate*:

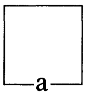
> x is an aggregate =df. x contains a pair of discrete parts that are not connected, mediately or immediately, by relations of specific dependence.

7. SUBSTANCE

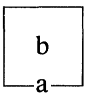
Our job, now, is to construct a diagrammatic language capable of representing perspicuously the above-mentioned categories of entities.

Substances are independent atoms, accidents are dependent atoms. An independent atom is an entity that is in need of no other individual entity outside itself in order to exist. Examples of substances might be: a thinker, a stone, a separated twist of DNA (if such there be).[17] Examples of substantial entities which are not substances might be: a hemisphere of a planet, a twist of DNA in Chisholm's leg.

To depict substances we shall employ solid frames, which are to be understood in some respects after the manner of the ovoids of Euler or Venn diagrams.[18] Thus:

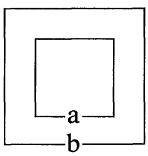


shall depict the independent atom whose proper name is a. We shall follow the convention that a proper name refers to the entire contents of the frame upon which it is inscribed. We shall adopt the (broadly Wittgensteinian) convention that all entities (whether individual or not) have a unique proper name and all proper names denote unique entities. Unlike Wittgenstein however we shall allow that proper names may depict parts of what other proper names depict. Thus for example



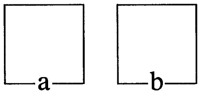
shall signify that b is a part of the independent atom a. (The diagram does not tell us whether b is or is not an atom, i.e. it does not tell us to what formal category the individual b belongs. Nor, incidentally, does it tell us whether b is a proper or improper part of the substance a.)

Because no substance is ever a proper part of another, it follows that



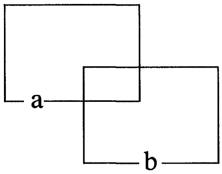
is always false.[19] Here we accept Aristotle's view: a substance does not have any substances as proper parts (though of course every substance can, in different ways, be dismembered into substances – for example into arms and legs, or conceivably also into separate molecules).

Given our definitions, it will turn out that



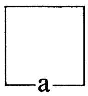
which signifies that independent atom a is discrete from independent atom b, is in contrast always true.

If one substance could overlap with another in the sense of sharing some individual part in common, without either including or being included in this other, then this state of affairs would be represented by:

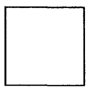


Siamese twins do not constitute an example of a structure such as this, since neither twin is, before separation, a substance. Touching billiard balls would not constitute an example, either, since the portion of boundary they share is not an individual, on our account.

Degen's law allows us to infer from

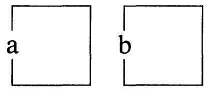


to:



The latter may for the moment be read as signifying that some independent atom exists.[20]

From

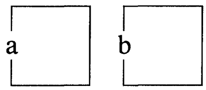


we may infer:

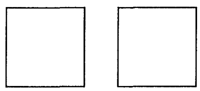
a b

The latter is, be it noted, a *fact* in the sense of TLP 3.142: it is not a list, or set, of names.

From

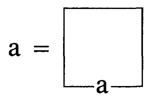


we may also however infer:



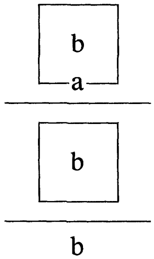
Read: ‘two independent atoms exist.’

The metadiagrammatic:



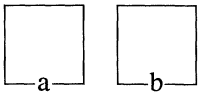
which comes close to saying what Wittgenstein says cannot be said at 4.126, is either necessarily true or necessarily false, according to whether a is or is not a substance.

We also have the inference:



i.e. from b is a part of independent atom a we may infer: b is a part of some independent atom, and from this we may infer: b exists.

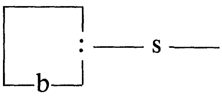
Note that from:



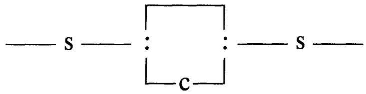
we cannot infer that a and b are disconnected, i.e. that they are not linked via relations of dependence through the mediation of other entities. The fact that an entity is a mere aggregate in this sense is a negative fact and therefore (as we have seen) not something that can be directly depicted. The property of being an aggregate is not diagrammatically perspicuous: not everything that is depicted by a non-connected diagram is itself non-connected in reality.

8. ACCIDENTS

An accident or dependent atom a is an entity which is necessarily such that it requires some other discrete entity or entities in order to exist but not vice versa. The required entities (here substances) are called the *carriers* or *termini* or *fundamenta* of the dependent atom.



shall depict a dependent atom b with one terminus.



shall depict a dependent relational atom c with two termini, etc.

The 's' inscribed on the links connecting dependent frames to their carriers here signifies the relation of specific dependence.

Thus a picture of a dependent atom comprises:

> (1) a frame with one or more perforated walls from each of which protrudes (2) a dependence link representing the specific dependence of the dependent atom upon the relevant carrier or fundament.

(1) and (2), here, are syncategorematic parts of the diagrams in which they occur. That is to say, neither is well-formed (categorematic) for the purposes of Degen's law inferences.

The diagrammatic equivalent of the notion of closure introduced above is the concept of a closed diagram, i.e. a diagram all of whose links are connected to the walls of other frames. This suggests the following

> **Closure Law**: If θ is a diagram in which dependent atoms or sub-atoms are depicted, then there is some unique smallest diagram cl(θ), the closure of θ, in which all the relevant containing atoms and independent carrier-atoms are depicted.[21]

It is possible that dependent atoms may themselves serve as carriers for further dependent accident-like entities of a higher order. For example the individual redness of my bruise is dependent on the bruise itself, which is in turn dependent on me. Such chains of one-sided dependence-relations must however come to an end after a finite number of steps. Dependent atoms never occur alone, but are in every case constituents of molecules in which their carriers are also contained. Thus we can embrace the following:

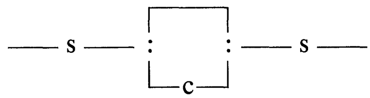
> **Principle of Ontological Well-Foundedness** (***ontologische Begründungsaxiom***)**:** That on which a dependent atom depends is always such as to include one or more independent atoms as parts.

This may be also be formulated as the:

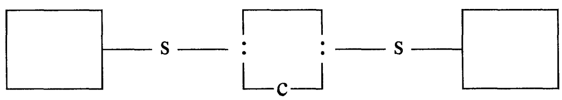
Strong Law of Immanent Realism: If there is anything, then there is a substance.

We would more precisely need to affirm a principle to the effect that there are a finite number of dependence steps between dependent atom and independent carrier – a principle to the effect that, leaving aside the mutual dependence of sub-atoms, every dependence diagram is a finite non-cyclical graph.

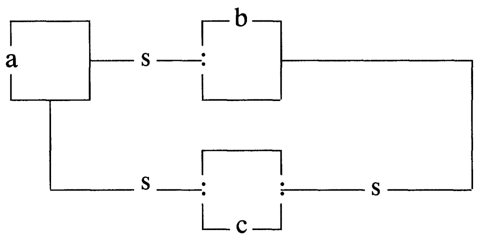
The principle of well-foundedness does not, however, allow us to infer e.g. from



to:



since we also have cases such as the following, where a relational accident involves only a single substance:



Here a is (for example) a thinker, b an act of seeing, c a state of enjoying b. (Aristotle, incidentally, affirms that accidents of accidents are always also accidents of substances in this sense.[22])

9. SUB-ATOMS (MUTUALLY DEPENDENT PARTS OF ATOMS)

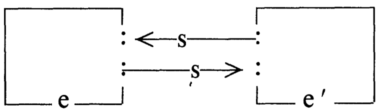
We have distinguished between essential parts of a substance, for example Rupert's *being heavy* or *having a shape*, and accidents of a substance, for example Rupert's present headache, or present jump. And we have drawn a similar distinction also in relation to accidents, for example between the brightness of a colour, on the one hand, and the noise of a walk, on the other. (No colour can exist without some brightness; a walk can however exist without it being the case that there is an accident of noise inhering in it.) It is for the moment of no consequence that the line between the two sorts of individual property may be somewhat difficult to draw. Our initial concern is merely to construct a directly depicting language that is able to cope with both sorts of case.

We shall assume in what follows that sub-atoms always stand in mutual dependence relations to each other. A picture of a sub-atom shall therefore consist of:

(1) a frame with one or more broken walls from each of which protrudes

(2) a dependence arrow representing the dependence of the dependent atom upon some one or more other sub-atoms and upon which is inscribed an 's' or 'b', indicating that the dependence involved is specific or boundary dependence, respectively.[23]

In the simplest case, sub-atoms stand in relations of pair-wise mutual specific dependence, for example as follows:

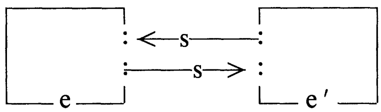


Clearly, similar diagrams can be constructed where n-fold mutual dependence obtains, and the idea is that the internal structure of every atom could be represented exhaustively by a family of complex diagrams of the given form, representing cuts of different sorts and in which all sub-atoms are eventually displayed.

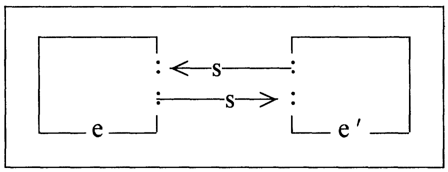
The arrow here indicates direction of dependence and serves to distinguish mutual dependence from the one-sided dependence involved in the case of dependent atoms. In all closed diagrams, each dependence

arrow connects its frame to the wall of some other frame. A picture of a sub-atom is further characterized by the fact that in a closed diagram it is reciprocally connected by dependence arrows to other frames.

From



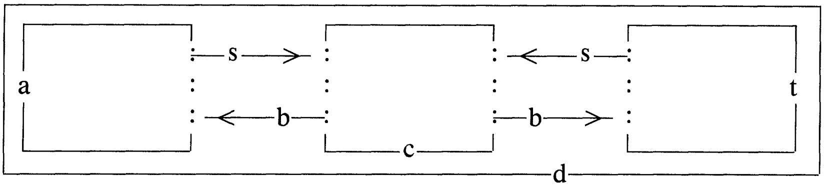
we may infer



For if e and e' really do depend upon nothing other than themselves in order to exist, then together they must constitute an independent atom.

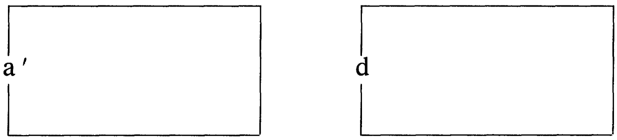
10. BOUNDARIES AND BOUNDARY DEPENDENCE

We are supposing, with Aristotle (though in a different terminology), that Darius is an independent atom. Then by definition of atom, no part of Darius is able to exist except in association with all the other parts. Let 'a' signify Darius's arm, as it is, now, attached to the remainder (the torso) t of Darius. Then there is some boundary c running between a and t, and inspection reveals that c is boundary dependent on both a and t, by our definition above. In particular, c is a part of both a and t. Boundaries are in this sense – and like universals – *multiply located.*[24] Moreover, both a and t are specifically dependent upon c. For a and t are in fact defined and delineated only via c. Any alternative delineation would capture not a and t but more or less distant cousins. a and t, then, exist as a matter of necessity only if c exists. But this is just the definition of specific dependence. The whole of Darius, on this particular parsing of his structure, may accordingly be represented thus:



Note that this representation is in order as it stands, in spite of the fact that c is part of both a and t, because it is not an individual part.

How, now, does Darius's arm a ′ relate to Darius, when once it has been actually removed from Darius's torso? Precisely thus:

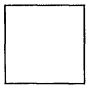


Note that we have here identity of Darius himself across the temporal interval during which the scission takes place. Darius, as a substance, is self-identical from the beginning to the end of his existence. But neither before nor after the removal of his arm is Darius identical with any substantial part of himself (something which follows trivially from the definition of substance and from our assumption that Darius is such as to satisfy this definition). After the operation Darius (now minus arm) is still an atom in his own right. The two diagrams allow us to see very clearly why this is so. a and a′ are non-identical, because a′ is a substance and a is merely substantial. No such difference of category arises in the case of Darius before and after the loss of his arm.

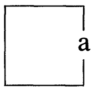
Boundaries correspond to possible cuts or parsings through reality. All extended objects allow an indefinite number of cuts or parsings of this sort, which are typically skew to each other. This possibility is, for Brentano, a mark of what is continuously extended.[25] Skew to *all* these parsings, however, are divisions, e.g., into sub-atoms.

11. UNIVERSALS

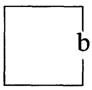
Return, now, to the diagram:



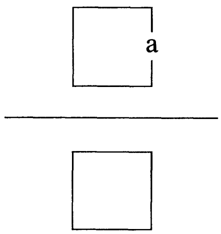
for which we suggested the reading 'some independent atom exists'. Can this diagram truly be counted as satisfying the conditions on a directly depicting language? To see why there is a problem here, consider the following experiment. Imagine a pair of distinct diagrams, say



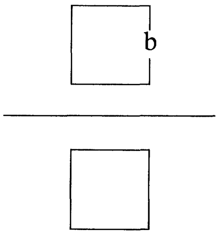
and



inscribed on separate sheets of paper. Suppose, further, that both diagrams are true, and in such a way that they depict discrete entities which we might assume to be several miles apart. Imagine, now, that Degen's law is applied to each diagram, so that on our separate sheets of paper we now have:



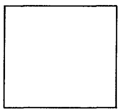
and:



Do the two resulting diagrams, now, depict identicals? This seems counterintuitive, given that the initial diagrams depicted discrete and separate entities, and that the two resulting diagrams are such as to depict parts of these entities. On the other hand it seems that some, at least, of our (diagrammatic) intuitions point in the opposite direction. These intuitions, which amount to the thesis that there can be no ambiguous diagrams – or as we expressed it in our treatment of Bergmann, above: every (true) diagram stands for one and only one existent – can be packaged together as the:

> **Principle of Diagrammatic Rigidity:** If θ depicts x, then there must be something in θ in virtue of which it does so.[26]

It follows from this principle that (true) equiform diagrams picture one and the same entity. But what can this entity be, in the case of diagrams like:

 ?

The only candidate which here presents itself would seem to be the formal universal *substance*, an entity which would serve as a non-individual part of every individual substance. If this is correct, then the diagram in question can be read as signifying either:

> Some individual substance exists,

or:

> The universal *substance* exists (is realized, is exemplified).

This argument (or ontological experiment) is, I believe, sound and yields the conclusion that certain formal universals exist. Moreover, it can be extended in such as way as to lend credence also to the thesis that there are material universals such as the *humanitas* of Rupert, or the *redness* that is shared in common by Rudolf's nose and this telephone box. The diagrammatic language can, accordingly, be extended to depict also material universals in this sense. Some classical positions as to the nature of universals then fall out as incapable of realization within the framework of a universal characteristic. Other such positions, however, can be so realized, in ways which yield alternative directly depicting languages of surprising force. The issue as to which of these alternatives comes closest to depicting the joints of reality remains, however, open.

NOTES

[1] Peirce 1933, 4.530. All references to Peirce in this section are taken from this passage.
[2] See Descartes' letter to Mersenne of 20 November 1629, Engl. trans. in Descartes, *Philosophical Letters*, A. Kenny, ed., Oxford: Blackwell, 1970, p. 6.
[3] Meyer 1957, p. 57.
[4] Quoted in Brentano 1982, pp. x-xi.
[5] Dummett 1981, p. 482.
[6] We recall that Gauss, too, conceived algebra as a 'science of the eye', a matter of observation, though of objects of a highly recondite character: cf. Peirce 1.34, 4.233.
[7] Cf. Peirce, 4.356.
[8] See e.g. Aristotle, *Met.*, 1027 b 22, 1051 b 32ff.
[9] See p. 149 of "Notes on Ontology".
[10] Kit Fine's 1985 rests essentially on a rejection of this idea, and makes possible for the first time a coherent formulation of an ontology in relation to which variables might properly be exploited in a directly depicting language.
[11] See Degen 1978, Smith and Mulligan 1982, pp. 81-91.
[12] Thus in particular Bergmann's notation of concentric circles cannot serve as the basis of a directly depicting language, since it allows disjunction to be treated as just another character (1981, p. 150).
[13] Much of this section bears a strong indebtedness to Chisholm's forthcoming work on the theory of categories. See also his 1982a, 1982b, 1984.
[14] 1981, p. 145.
[15] 'Part', here, means always 'proper or improper part'.
[16] Cf. *Tractatus*, 3.328, 5.47321.
[17] In some ontologies the atoms are constituted not by substances (ordinary objects) but by other sorts of entities. The differences here may be material – as in the case where one takes sense data as atoms. Or they may be formal: thus *Sachverhalte*, on one reading of the *Tractatus*, take the place of our atoms, and Wittgensteinian simple objects take the place of our sub-atoms.

[18] See Smith and Mulligan 1982 and 1983 for more details of this aspect of the theory.
[19] 'Always' in the sense that it is false for every pair of proper names 'a' and 'b'. (Empty proper names have, it will be remembered, been excluded from consideration.)
[20] It is something like a prototype in the sense of *Tractatus*, 3.333.
[21] It is worth noting that 'cl( )' here satisfies the usual Kuratowski axioms for a topological space. It is worth noting also that, like the converse Degen laws considered above, the closure law is not effective in the sense that we are not, in general, able to construct the diagram whose existence the law would purport to guarantee.
[22] For a discussion of this law in Aristotle and Brentano see my 1987.
[23] Again, neither (1) nor (2) is when taken alone well-formed for the purposes of Degen's law inferences.
[24] There is, in fact, as Brentano acknowledged (1976), a deep and surprising parallel between boundaries and universals, so that much of what is necessary for an adequate diagrammatic treatment of the latter can be derived from the treatment of the former.
[25] See, again, his 1976.
[26] Cf. also Bergmann's 'fundamental principle of ontology': different complexes must differ in a constituent. (Cf. his 1967, p. 22.)

Note that the principle given in the text runs counter, e.g. to Wittgenstein's idea that the depiction of an object or complex might be catered for in part psychologically, in terms of some extra-diagrammatic 'projection', at 3.11-13. The psychological connotations are incidentally no longer present at 4.0141.

REFERENCES

Bergmann, Gustav 1967 *Realism. A Critique of Brentano and Meinong*, Madison: University of Wisconsin Press.

______, 1981 "Notes on Ontology", *Noûs*, 15, 131-64.

Brentano, Franz 1976 *Philosophische Untersuchungen zu Raum, Zeit und Kontinuum*, ed. by R. M. Chisholm and S. Körner, Hamburg: Meiner, Eng. trans., *Philosophical Investigations on Space, Time and the Continuum*, by B. Smith, London: Croom Helm, 1988.

______,1982 *Deskriptive Psychologie*, R. M. Chisholm and W. Baumgartner, eds., Hamburg: Meiner.

Chisholm, R. M. 1982 *Brentano and Meinong Studies*, Amsterdam: Rodopi.

______, 1982a "Beginnings and Endings", in Chisholm 1982, 114-24.

______, 1982b "Brentano's Theory of Substance and Accident", in Chisholm 1982, 3-16.

______, 1984 "Boundaries as Dependent Particulars", *Grazer Philosophische Studien*, 10, 87-95.

Degen, Wolfgang 1978 "Sketch of a Rational Grammar", unpublished MS., University of Erlangen.

Dummett, M. A. E. 1981 *The Interpretation of Frege's Philosophy*, London: Duckworth.

Fine, K. 1985 *Reasoning with Arbitrary Objects*, Oxford: Blackwell.

Grossmann, R. 1983 *The Categorial Structure of the World*, Bloomington: Indiana University Press.

Husserl, E. 1900/01 *Logische Untersuchungen*, 1st ed., Halle: Niemeyer, 2nd ed., 1913/21; Eng. trans. by J.N. Findlay, *Logical Investigations*, London: Routledge and Kegan Paul, 1970.

Meyer, R. W. 1957 "Joachim Jungius und die Philosophie seiner Zeit" in *Die Entfaltung der Wissenschaft*, Hamburg: J. J. Augustin.

Peirce, C. S. 1933 *Collected Papers*, Cambridge, Mass.: Harvard University Press.

Smith, B. 1987 "The Substance of Brentano's Ontology", *Topoi*, 6/1, 39-49.

Smith, B., ed. 1982 *Parts and Moments. Studies in Logic and Formal Ontology*, Munich: Philosophia.

Smith, B. and Mulligan, K. 1982 "Pieces of a Theory", in Smith, ed., 15-109.

______, 1983 "Framework for Formal Ontology", *Topoi*, 3, 73-85.

KAREL LAMBERT

DEFINITE DESCRIPTIONS AND THE THEORY OF OBJECTS*

1. A NEW EXPLANATION

The purpose of this section is to provide a new explanation of the various approaches to the treatment of definite descriptions. What is to be called *the naive theory of definite descriptions* has not existed prior to this study. But the explanation is not any less illuminating for that fact. What this has to do with the theory of objects will emerge in due course.

Assume, then, a first order language consisting of predicates, variables, perhaps simple singular terms, truthfunctional connectives, quantifiers and identity. Let the underlying logic be classical first order logic with identity. Without simple singular terms the language and logic so outlined is essentially the canonical idiom of Quine's *Word and Object*.[1] For that reason the *full* language and logic is here denominated Q^+.

To Q^+ let there be added the operator symbol 'I'. When prefixed with a variable to an open sentence it produces a (complex) singular term – for instance, it produces an expression of the form '$\mathrm{I}x(Px)$', where P is a one place predicate of Q^+. Such an expression, called a *definite description*, is to be read as 'the object (x such that x is (a)) P'.

Consider now the following question:

(1) What object does the definite description '$\mathrm{I}(Px)$' refer to?

A very natural answer is:

(2) It refers to whatever object is such that it and it only is (a) P.

By semantic descent (2) yields:

(3) For any given object, it is $\mathrm{I}x(Px)$ if and only if it and it only is (a) P.

This sentence, in turn, is a very natural answer to the question:

(4) What object is $\mathrm{I}x(Px)$?

Generalizing (3) to any definite description, the result of adding the axiom scheme

$$A_1 \quad \forall x(x = \mathrm{I}yA \equiv\cdot\ (Ax/y\ \&\ \ \forall y(A \supset x = y)))$$

to Q^+ is *the naive theory of definite descriptions* (*NTDD*). (In A_1, A is a sentence and At/s is the result of replacing all occurrences of the singular term or (free) variable s in A by the singular term or variable t.)[2]

Consider now the following three important statements.

$$T_1 \ A\mathrm{I}yA$$
$$T_2 \quad \forall x(x \epsilon \hat{y}A \equiv A)$$
$$T_3 \ \mathrm{I}y(y \neq y) = \mathrm{I}y(y \neq y) \equiv \mathrm{I}y(y \neq y) \neq \mathrm{I}y(y \neq y)$$

T_1, which is derivable from A_1 in *NTDD*, is the paradoxical Meinongian principle Russell exploited in his famous 1905 "demolition" of Meinong's theory of objects.[3] For despite its naturalness, it yields the contradictory sentence

(5) $\quad P\mathrm{I}y(Py\ \&\ {\sim}Py)\ \&\ {\sim}\ P\mathrm{I}y(Py\ \&\ {\sim}Py).$

in *NTDD*.

T_2, which is derivable in *NTDD* (when the vocabulary of *NTDD* is supplemented by 'ϵ') via the very natural definition,

$$D_1\ \hat{y}A \ =\mathrm{df}\ \mathrm{I}y\ \forall x(x \epsilon y \equiv A),$$

is a version of the paradoxical principle of comprehension in naive set theory. It yields, as Russell first showed, the imminently contradictory sentence

(6) $\quad \hat{y}(y \notin y) \in y(y \notin y) \equiv \hat{y}(y \notin y) \notin \hat{y}(y \notin y)$

which bears his name.

Apparently Russell did not think the paradoxical consequences, T_1 and T_2, were related. But in fact they are; they have a common source, A_1 of *NTDD*. The point is important vis à vis *NTDD* because Russell's attitude was much milder in the case of the paradoxical principle T_2 than in the case of the paradoxical principle T_1; he did not think the odious consequences of T_2 "demolished" set theory but he did think the equally odious consequences of T_1 "demolished" Meinong's theory of objects.

T_3, which is structurally similar to (6), is rather direct evidence of the paradoxical character of A_1 itself. This sentence may be called the paradox of the nonselfidentical object–or, if you prefer, the paradox of the egoless object.

For the moment fix on the paradoxical principle of comprehension, T_2. It is the principle of naive set theory to which various contemporary approaches to sets are reactions.

One such approach *amends* A_1. For instance, Quine in the spirit of Zermelo replaced T_2 by the following variant of the *Aussonderung* principle in his system *ML*:

$$(7) \qquad \forall x(x \epsilon \hat{y} A \equiv (x \epsilon V \,\&\, A)).$$

(Zermelo did not recognize a universal class V nor did he have abstracts.) This principle does not yield the imminently contradictory Russell sentence. It yields only the conclusion that the Russell set – $\hat{y}(y \notin y)$ – is not a member of the universal set V.

A second approach to the problems generated by T_2 retains T_2 but *restricts the formation rules of naive set theory*. Russell is the outstanding exponent of this approach, an approach embodied in his version of type theory. The offensive expression '$(y \notin y)$' is no longer regarded as a logically grammatical open sentence. So the Russell sentence (6) is not even a candidate for inference from T_2.

A third approach restricts not the formation rules of naive set theory but rather *its substitution rules*. This is the approach embodied in *Quine's New Foundations*.[4] Though grammatical, '$(y \notin y)$' cannot be substituted into A in T_2 because A can only be replaced by stratified sentences, a property that the sentence '$(y \notin y)$' lacks.

A final approach leaves T_2, the syntax of naive set theory and its substitution rules in tact. Instead it *rejects the underlying logic*. This approach is embodied in different ways in intuitionistic set theory and free set theory. The former blocks the inference from the Russell sentence (6) to contradiction by rejecting the principle of excluded middle; the latter blocks the inference from T_2 to the Russell sentence by disallowing instead the use of the principle of universal specification *except* when there is something that is the specified object. In free set theory, T_2 yields only the conclusion that there is no such set as the Russell set.

Now return to A_1. It is the distinctive axiom of the naive theory of definite descriptions. The various approaches to the treatment of definite descriptions can be understood as reactions to A_1 just as various approaches to sets can be seen as reactions to the (naive) principle of comprehension.

Consider, first, Frege's theory of definite descriptions. This theory is obtained by replacing A_1 with the following axiom scheme:

$$A_2 \; x(x = \mathrm{I}yA \equiv ((\mathrm{A}x/y \;\&\; \forall\, y(A \supset y = x)).\ \mathrm{v}.\ (\sim \exists\, z\ \forall\, y\ (A \equiv y = z)\ \&\ x = *)),$$

where '$*$' is either an arbitrary chosen object in one version of the theory or $\hat{y}A$ in the other version of the theory. No matter which formulation is used, however, it is clear that Frege's approach is an example of the amendment approach. This reaction to *NTDD* allows the inference from A_2 to T_1 and T_2 only under the special conditions that $\mathrm{I}yA \neq *$ and $\mathrm{I}y(\forall x(x \epsilon y \equiv A)) \neq *$ and the inference to T_3 on the condition that $\mathrm{I}y(y \neq y) \neq *$.

Consider next the Hilbert-Bernays theory of definite descriptions. It restricts the formation rules of *NTDD*. An expression of the form '$\mathrm{I}yA$' is grammatical only on the condition that the uniqueness of A is *provable*. Hence the offensive instances of T_1 and T_2 are not derivable. Nor is T_3; all such instances are banished to the realm of logically ungrammatical expressions.

Next is Russell's theory of definite descriptions – at least as that theory is expressed in *Principia Mathematica* and in the *Introduction to Mathematical Philosophy*. Russell apparently does not declare ungrammatical expressions of the form '$\mathrm{I}yA$' when A is not unique. Rather the substitution rules are restricted to *singular terms* only in the places of the free variables in the valid sentences. Since expressions of the form '$\mathrm{I}yA$' are not genuine singular terms for Russell, application of the principle of universal specification to definite descriptions is disallowed and the inferences to T_1 through T_3 are successfully blocked. As mentioned earlier, T_2 is asserted independently and contradiction is presumably avoided by restricting the syntax of naive set theory. One sees, therefore, an important difference between Russell's treatment of definite descriptions, which are not ungrammatical short of set theory, and his treatment of sets – he would have said "classes". Of course, Russell also provides well known elimination rules for contexts of the form $\mathrm{BI}yA/x$ and $\mathrm{E!I}yA$ in the early definitions of chapter *14 in *Principia Mathematica*. But these are not central to the explanation of the origins of definite description theories being developed here.

The last approach, the free description theory approach, originated initially with the writer, and soon afterward, other versions were produced by Rolf Schock, Bas van Fraassen, David Kaplan, and Dana Scott, among others. Like intuitionistic and free set theory, free description theory rejects the logic underlying *NTDD*. For example, it

allows application of the principle of universal specification to sentences containing definite descriptions only on the condition that there is the object so described. As a result A_1 is retainable in its original form but T_1 through T_3 are derivable only on the condition a statement of the form '$\exists\ x(x = \mathrm{I}yA)$' is true. In particular, T_2 is independent of, but compatible with, the free theory of definite descriptions, given, that is, that set abstracts are treated as in the definition D_1.

This completes the new explanation of the various approaches to the treatment of definite descriptions. To sum up, they are reactions to the very natural principle A_1 which, however, is paradoxical in the context of classical first order logic with identity. Moreover the ways out of paradox in the case the theory of definite descriptions bear a striking resemblance to certain ways out of paradox in the case of set theory.

2. AN APPLICATION OF THE FOREGOING EXPLANATION

It is not news that a certain Polish tradition in logic deriving from Lesniewski conceived the classical logic of predicates as a litany of the most general laws of being. Hence the name "Ontology" for that segment of logic. Similarly, what I want to suggest now is that *NTDD* and its progeny be interpreted instead as theories of objects à la Meinong. Such an interpretation, indeed, is latent in the reading of an expression of the form '$\mathrm{I}y(Py)$' as 'the *object* (x such that x is (a)) P'. Of course this means that the quantifiers must be dissociated from their normal existential force. Consider it done forthwith.

If woefully incomplete – for example, there are no axioms about incomplete objects or even for special complete objects such as those Meinong call *Objektive* – the various theories of definite desciptions so interpreted at least lay out certain basic rudiments of object theory. Moreover, even these rudimentary theories are sufficient to the discussion of certain fundamental problems in object theory – for example, the problem whether the object which both is a spheroid and is such that it is not a spheroid (herafter $\mathrm{I}y(Sy\ \&\ {\sim}\mathrm{S}y)$ is a kind of nonexistent object or no object at all. This is no small matter for object theory. Russell once called for courage when confronted by $\mathrm{I}y(Sy\ \&\ {\sim}Sy)$ because one cannot be clear about the class of supposed nonexistent objects unless one has a theory encompassing the "object" in question.[5]

In the spirit of interpreting theories of definite descriptions as theories of objects, *NTDD* may be construed as the (*rudimentary naive theory of*

(*complete*) *objects, NTCO*. As such, it invites comparison with a fragment of Meinong's theory of (complete) *Objekte*. One reason, as pointed out earlier, is that it suffers from the same faults; that is, it yields the paradoxical Meinong-like principle T_1.[6] Another favorable but nonderisive reason is that it contains the theorems

(8) $\exists x(x = \mathrm{I}yA)$,

which says, in effect, that IyA is an object, and

(9) $\forall x \exists y(x = y)$;

that is, everything is an object.

Returning to D_1, that definition is testimony to the basic character of object theory as conceived here; it shows that sets are simply certain kinds of objects. So insofar as mathematics emerges from set theory, it would seem, therefore, that an even more basic theory of mathematical interest is the theory of objects. Similarly, properties can be construed as another kind of object, at least when modal operators and the relation of possession are added to the basic language. Thus the property (of being) so and so can be construed as the object such that everything necessarily possesses it if and only if it is so and so. In symbols this is:

$$D_2 \ \lambda yA =_{\mathrm{df}} \mathrm{I}y \ \forall \ \mathrm{x}(\text{necessarily } (\mathrm{x} \text{ possesses } \mathrm{y} \equiv \mathrm{A})).$$

And the same holds for other kinds of objects perhaps in the manner of Quine's reduction of all singular terms to definite descriptions.

Since *NTCO* is unsound, the Fregeian, Hilbert-Bernays, Russellian and free object theories can be seen as attempts to avoid or remove the features of the naive theory of objects – that made it unsound – such as Russell's famous discovery vis à vis T_1. In this respect, they succeed when attention is limited to the elementary language with identity and definite descriptions; they are each of them provably sound.

How, then, are they to be judged? One way to assess them is via their set theoretical consequences *given* the definition D_1 of sets as special kinds of objects.

Consider first Frege's theory – or rather the theory of objects constructed in the image of Frege's chosen object theory of definite descriptions where the chosen object is the so called nonexistent entity.[7] This theory of objects treats every definite description as standing for a genuine object, even Iy(Sy & $\sim Sy$). The latter, and any other definite description referring to *, do not, of course, refer to existents. However

the Frege-inspired theory of objects has an unpleasant result vis a vis those objects that are sets; it violates the principle of set extensionality in the presence of D_1.

To see this notice first that this object theory has the consequence that

$$FT_2\ \hat{y}A \neq * \ \forall x(x \in \hat{y}A \equiv A).$$

But this yields both

$$(10) \qquad \hat{y}(y \notin y \lor y = 0) = *$$

and

$$(11) \qquad \hat{y}(y \notin y \lor y = 1) = *$$

Hence

$$(12) \qquad \hat{y}(y \notin y \lor y = 0) = \hat{y}(y \notin y \lor y = 1)$$

by the transitivity of identity. But this result conflicts with the principle of set extensionality because it is not true that

$$(13) \qquad \forall x(x \in \hat{y}(y \notin y \lor y = 0) \equiv x \in \hat{y}(y \notin y \lor y = 1)).$$

Turning next to the Hilbert-Bernays like theory of objects, it, too, holds that every definite description with a grammatically sound basis picks out an object. But since no definite description whose basis is not provably unique is even grammatical the question whether $\mathrm{I}y(Sy\ \&\ {\sim}Sy)$ and $\hat{y}(y \notin y)$ are objects simply does not arise. The cost is even higher when one realizes that since provability is not decidable, it is not always determinable whether the question of objecthood arises. Moreover, even if it be granted that the uniqueness of the bases of the definite descriptions '$\mathrm{I}y$(y is an author of PM)' and '$\mathrm{I}y$(y is a planet causing the perturbations in Mercury's orbit)' are not provable, it seems alien to object theory not to recognize these at least as nonexistent objects.

In a theory of objects in the Russellian mode,

$$\exists\ x(x = \hat{y}A) \supset \forall\ x(x \in \hat{y}A \equiv A)$$

is derivable. Unlike the Frege conditioned theory of objects in which both $\hat{y}(y \notin y)$ and $\mathrm{I}y(Sy\ \&\ {\sim}Sy)$ qualify as objects, and unlike the Hilbert-Bernays theory in which their objecthood does not arise, in the Russell-like object theory neither qualifies as an object. Of course the existent objects now form a subset of the objects, but $\hat{y}(y \notin y)$ and $\mathrm{I}y(Sy\ \&\ {\sim}Sy)$ are not merely nonexistent objects, they are not objects at all. It must be

remembered that in the Russellian object theory the nonobjecthood of the two items in question does not conflict with classical logic because of Russell's stance that definite descriptions are not singular terms. But like the Fregeian theory, this theory also conflicts with the principle of set extensionality; the Russell elimination rules yield the verdict that $\hat{y}(y \notin y) \neq \hat{y}(y \notin y)$, a verdict which clearly conflicts with the principle in question.

Finally, it is time to consider the object theory suggested by free description theory. This theory is simply the naive theory of objects sans the underlying classical logic. Like the Russellesque theory of objects, it must deny objecthood both to $\hat{y}(y \notin y)$ and $\mathrm{I}y(Sy \;\&\; {\sim}Sy)$, but since '$\hat{y}(y \notin y) = y(y \notin y)$' is a simple substitution instance of the reflexivity of identity, there is no conflict on this score with the principle of set extensionality. Nor is there conflict with that principle in any other regard. Moreover free object theory is also compatible with the view that the detective who lived at 221 B Baker Street is an object though a nonexistent one but that $\mathrm{I}y(Sy \;\&\; {\sim}Sy)$ is neither an existent nor a nonexistent object because it is no object at all.

To sum up, the major problem of the Fregeian conditioned object theory is that though it treats $\hat{y}(y \notin y)$ as a genuine but nonexistent object, some way has to be found for identifying objects in general which does not conflict with set extensionality. The Hilbert-Bernays theory is hopelessly restrictive. The Russellesque theory must also find a way to cope both with the nonobjecthood of $\hat{y}(y \notin y)$ and $\mathrm{I}y(Sy \;\&\; {\sim}Sy)$ and the principle of set extensionality. Finally the free object theory has similar problems to the Russellesque object theory except for the problem of set extensionality.

Let me finish with a suggestion about how to deal with the nonobjecthood of $\hat{y}(y \notin y)$ and $\mathrm{I}y(Sy \;\&\; {\sim}Sy)$. In the first place, it should be noted that not all object theorists accept the objecthood of $\mathrm{I}y(Sy \;\&\; {\sim}Sy)$. A case in point is Terence Parsons whose own development in his Meinong inspired book *Nonexistent Objects* countenances all sorts of nonexistent objects. In the second place, cases like $\hat{y}(y \notin y)$ seem to fall in a category close to what Meinong called *defective objects*. One way to identify this category may be syntactical; that is, a defective object is not an object per se but only a loose, picturesque way of talking about various kinds of ill formed set abstracts (of definite descriptions). Then type theory, for example, can be used to isolate the defective object which is $\hat{y}(y \notin y)$ – the Russell set. This would be the direction of Russell's own

development in *Principia* in which the principle of comprehension is asserted in the context of type theory. But there is another nonsyntactical way to view defective objects, and that is the way exemplified in Scott's approach to free set theory.[8] In this approach unrestricted set abstraction is added to free object theory sans type theory. Since only a restricted principle universal specification holds, one can prove only the nonobjecthood of $\hat{y}(y \notin y)$. Such sets are examples of what Scott calls *virtual objects*, but as sets they must obey basic principles like set extensionality. Yet virtual objects aren't objects any more than a blunderbus is a bus. One nice thing about this suggestion is that it can also accommodate $\mathrm{I}y(Sy \;\&\; {\sim} Sy)$; it can also be construed as a virtual object, though, of course, not of the same sort as $\hat{y}(y \notin y)$. Because Scott disallows quantification over virtuals, the Meinongian principle that everything is an object, is not thereby violated.

NOTES

* I am grateful to the other participants in the conference on Formal Ontology Today for their various suggestions during the discussion of these remarks. But special thanks must be extended to five persons; first, to Penelope Maddy for suggestions, corrections, and her passion to keep me on the straight and narrow; second, to Kevin Mulligan who encouraged the presentation of this quasi-technical highly speculative venture in its current form; third, to Barry Smith who proved that another version of the Frege theory of definite descriptions I presented at the conference was not equivalent to the more familiar version in this essay; and fourth, to Peter Simons and Graeme Forbes who made interesting remarks about what must be done to advance my speculations on object theory beyond the tickle-one's-fancy-stage. Finally, I am grateful to the National Endowment for the Humanities for the award of a stipend during the summer of 1988, a stipend which underwrote most of the research in this essay.

[1] W.V. Quine, *Word and Object*, New York (1960). See especially the chapter on Regimentation (Chapter 5).

[2] Among free logicians A_1 is known as Lambert's Law.

[3] See his "On Denoting," *Mind*, (1905), pp. 479-93.

[4] *American Mathematical Monthly*, 44 (1937), p. 70-80.

[5] See Russell's, "Review of A Meinong's *Untersuchungen zur Gegenstandstheorie und Psychologie", Mind*, NS14 (1905), p. 537.

[6] The use of the adjective "derisive" is intended in the sense exemplified by the mother who says derisively to her son, "You're just like your father! You'll never amount to anything!".

[7] The nonchosen object version isn't available because in this version not all set abstracts can be eliminated. So not all sets could be construed definitionally as objects of a certain sort.

[8] See his Existence and Description in Formal Logic" in *Bertrand Russell: Philosopher of the Century* (Ed. R. Schoneman). Allen & Unwin, London (1967).

HERBERT HOCHBERG

TRUTH MAKERS, TRUTH PREDICATES, AND TRUTH TYPES

In "The Concept of Truth in Formalized Languages," Tarski set out an equivalence condition for the introduction of a truth predicate into a linguistic schema.[1] Known as Convention-T, this condition requires that a "formally correct definition" of a truth predicate introduced into or defined in a metalanguage for a schema L, will be "adequate" if, for any statement of L, it has as a consequence a biconditional holding between a metalinguistic transcription of the statement and a metalinguistic sentence ascribing the truth predicate to the statement.[2] In his early paper, as well as in the later "The Semantic Conception of Truth," Tarski linked Convention-T to both "the classical Aristotelian conception of truth" expressed in Aristotle's assertion:

> To say of what is that it is not, or of what is not that it is,
> is false, while to say of what is that it is, or of what is not
> that it is not, is true.[3]

as well as to the notion that the truth of a sentence is determined by its correspondence to reality:

> If we wished to adapt ourselves to modern philosophical terminology, we could perhaps express this conception by means of the familiar formula:
> *The truth of a sentence consists in its agreement with (or correspondence to) reality.*
> (For a theory of truth which is based upon the latter formulation the term "correspondence theory" has been suggested.)[4]

As far as Tarski was concerned, "the pragmatic conception" and "the coherence theory" of truth have not "been put so far in an intelligible and unequivocal form."[5]

This paper will argue that Convention-T provides a ground for rejecting coherence theories as well as a ground for accepting a correspondence theory of truth that is in keeping with Aristotle's concise, but explicit, realism. The arguments will involve explicating the concepts of *correspondence* and *truth condition* in a way that construes a theory of truth quite differently than Tarski does. The discussion, in turn, will provide a basis for critically commenting on some recent proposed alternatives to theories of types. These alternatives attempt to use truth value gaps or related devices to deal with liar type paradoxes while avoiding construing the concept of truth as subject to the systematic ambiguity imposed by a theory of types. Finally, to meet some problems generated by the correspondence theory developed, we will briefly sketch an alternative correspondence theory.

To see why coherence theories cannot satisfy Convention-T, let 'aRb' be an atomic sentence and let '[aRb]' *designate* the proposition, where we can construe the proposition as either the sentence or as what is "expressed by" or "correlated with" or "meant by" the sentence. Thus, we can consider 'aRb' to belong to a basic language L and '[aRb]' to belong to a metalanguage for (or a suitably stratified extension of) L, L*. For the present we will confine the contexts to atomic sentences in order to facilitate the setting out of a version of the correspondence theory and to postpone consideration of questions about truth conditions for logically complex propositions.

We will take a coherence theory to hold that [aRb] is true if and only if there is some class of propositions C, such that [aRb] *coheres with* C. This involves recognizing a coherence relation obtaining between [aRb] and the class C (or a sub-set of C, or a conjunction of members of C, etc.). Then, according to a coherence theory, with 'Tr' as a suitable truth predicate, we have:

(D1) Tr[aRb] iff [aRb] *coheres with C.*

We need not further specify the relation *coheres with* in any way, for we can take it to be as strong as *entails* or as weak as *is consistent with.* However it is taken, (D1) will not allow us to derive the biconditional required by Convention-T, i. e. the T-sentence:

(E) Tr[aRb] iff aRb.

If we add C (the members of C) to the assumptions or axioms of the schema L* (or take C to consist of logical truths), then (E) may be

obtained, with *coheres with* taken as *is entailed by*. But, doing this would raise an obvious and familiar problem concerning C. For, if the members of C are added to the assumptions, then we take C (the members of C) to *be true* in a different and basic sense. Pointing out that the members of C are Tr in that each one is entailed by C is besides the point. If *coheres with* is not taken as *entails* but as basic or simply as *is consistent* with (or *probabilifies*, in some sense), (E) is not obtainable on the basis of (D1). To obtain (E), without taking *coheres with* in terms of entailment, one would have to assume

(A1) [aRb] *coheres with* C iff aRb

(or something equivalent or stronger). One might think that (A1) is trivial, since 'Tr' is defined in terms of 'coheres with'. One way of seeing the significance of Convention-T, for this discussion, is to realize that (A1) is far from trivial and does not follow from (D1).

Tarski's statement about a correspondence theory suggests expressing such a theory, for atomic propositions, in terms of a generalized version of (E):

(D2) Tr[...] iff ...,

where any atomic sentence may replace both occurrences of the dots. But, (D2) will only express such a theory, for atomic sentences, if we understand that the occurrence of the same sentence, within and without brackets, is used to express the connection between an atomic proposition (sentence) and a situation or possibility (state of affairs). When such a situation exists (obtains, is actual, is a fact), its existence furnishes the *truth ground* for the sentence on the right side of the biconditional and, hence, by (D2), for the left side of the biconditional as well.

Convention-T is then trivially satisfied by [aRb], as well as by every atomic sentence, given the adoption of the schema (D2) for the atomic sentences of L, *with the interpretation* we have just given to that pattern. We might note that Tarski's well known admonition that Convention-T does not furnish a definition of a truth predicate, but, rather, a criterion for determining the viability of purported definitions, is not germane here. His type of definition in terms of 'satisfaction', designed to deal with the evaluation of ascriptions of a truth predicate to non-atomic sentences of a formal schema, is trivially fulfilled by an atomic sentence which is satisfied (vacuously) by all objects of the domain if and only if its metalinguistic transcription holds. One could, quite trivially, make use of

an intermediary satisfaction relation.

The interpretation we have given to (D2) involves more than a general pattern for a T-sentence. We can make this point more explicit by replacing (D2) by the schema:

(D3) Tr[...] iff ([...] *represents* ... and ...).

In instances of (D3), such as:

Tr[aRb] iff ([aRb] *represents* aRb and aRb)

the sentence 'aRb' is understood to represent a situation, or possibility, when it occurs without brackets, while the complex sign consisting of the sentence within brackets is taken to stand for a proposition (or to be a metalinguistic sign designating the sentence). The first clause of the conjunction reflects our taking an atomic proposition to represent a situation, while the second clause is used to assert that the represented situation obtains or exists. As Wittgenstein concisely put matters in the *Tractatus*, a proposition "shows" how things stand if it is true and "says" that they so stand. What we can take to be shown is stated by the first conjunct in (D3), while what is said is stated in the second. But, taking such things to be stated relies on our understanding that atomic sentences represent situations or possibilities, which may or may not obtain.[6]

In place of a schema like (D3), with the understood role of the brackets and dots, one could express the view by using variables ranging over suitable entities, as in:

Tr(x) iff (∃ p)(x represents p & p).

Here 'p' ranges over atomic situations, both obtaining and non-obtaining, while 'x' ranges over atomic sentences or propositions. Moreover, given the reading of '(∃ p)(x represents p & p)' as 'there is a situation represented by x and that situation obtains', one would read '(∃ p)p' as 'there is a situation that obtains', just as in the case of an existential generalization from 'aRb' in '[aRb] *represents* aRb and aRb'. We shall return to this double use of atomic sentences, to represent situations and to assert they obtain, and the related use of bound variables, as in '(∃ p)p', and of 'obtains'.

Using (D3) we can derive (E). Clearly, we have 'Tr[aRb] → aRb'; and, understanding that

[aRb] *represents* aRb

is a theorem (or logical truth or linguistic truth or true by the semantical rules) of a suitable schema, we also have 'aRb $\rightarrow$ Tr[aRb]'. This was implicit in the use of the same sentence within and without brackets in (D2). The price one pays, on such a view, is the recognition of situations that may or may not obtain – or of potential and actual facts – or of existent and non-existent facts – or of facts which have being and which have no being. Thus, one takes quite literally the Aristotelian pronouncement that we speak truly when we say "of what is not that it is not" and Wittgenstein's statements in the *Tractatus*:

2.06 The existence and non-existence of states of affairs is reality. (We also call the existence of states of affairs a positive fact, and their non-existence a negative fact.)

4.1 Propositions represent the existence and non-existence of states of affairs.[7]

The points just considered bring out two themes that are implicit in a correspondence theory of the kind attributed to Aristotle and which Russell, Moore, and Wittgenstein revived early in this century, in setting forth a realistic theory of truth in opposition to the idealistic coherence theories of F. H. Bradley, B. Bosanquet, and H. H. Joachim. First, given that (D2) and (D3), properly understood, are really different ways of expressing the same view, such a theory holds that atomic sentences represent situations, which may obtain or not. This illustrates that a theory of truth, like the correspondence theory, is not stated simply by means of definitional patterns for a truth predicate. Rather, the correspondence theory is expressed by means of statements made *about* a pattern like (D3) and, consequently, *about* the interpretation of the atomic sentences of a purportedly perspicuous schema. Second, the type of correspondence theory we are considering, a theory of the kind held by Wittgenstein, Moore, and Russell, makes use of two senses of the term 'correspond', and this involves such a theory in the recognition of possibilities or potentialities.

In one sense, an atomic sentence (proposition) may be said to correspond to a situation irrespective of its truth or falsity, and, hence, irrespective of whether the situation obtains or exists. In a second sense, an atomic sentence is said to correspond to a fact, which is to say a situation that obtains or exists. The existence of the fact thus furnishes the condition of truth for the sentence. This double sense of 'corresponds' is

what Wittgenstein expressed by his use of 'shows' and 'says', what Russell expressed by speaking of positive and negative facts in "On Propositions: What They Are and How They Mean", and what Moore implicitly adhered to by speaking of existent and non-existent facts (and of facts that have being or have no being) in *Some Main Problems of Philosophy*. It also points to an ambiguity in the notion of a *truth condition* on a correspondence theory. In one sense a condition of truth is *the situation* that the atomic sentence or proposition represents, irrespective of the latter's truth value. In another sense the truth condition is *the obtaining* of such a situation, the existence of a fact. To avoid the inherent ambiguity involved, we can use the suggestive phrase 'truth ground' or 'truth maker' for this latter sense of 'truth condition'.

The point can be emphasized by considering a construal of 'Tr' that more literally reflects a common way of expressing a correspondence theory:

(D4) Tr[...] iff (∃ p)([...] corresponds to p)

where the variable 'p' now ranges over existent situations or *facts*. However, (D4) does not appear to allow us to derive a T-sentence unless we assume:

(A2) (∃ p)([...] corresponds to p) iff ... ,

or something similar. Thus, such a version of the correspondence theory seems to face the same problem raised in connection with the coherence theory. We shall return to this question when we consider an alternative version of a correspondence theory, without situations, at the end of the paper.

In addition to separating the two senses of 'correspond', a correspondence theorist must recognize that in order to derive the requisite T-sentences on a correspondence theory of truth, two kinds of statements must be employed in setting forth the theory. One kind involves the use of patterns like (D3), which explicitly contain a truth predicate, while the other provides the interpretation for such a pattern and specifies the ontological ground for taking the truth value, *true*, to apply to an atomic proposition or sentence. The ontological ground is an atomic fact, which is the truth maker for an atomic sentence. With (D2) this second component of the correspondence theory, the taking of facts as grounds of truth, is incorporated into the construal of the bracketing device and the interpretation of atomic sentences as representing situations. With

(D3) the correlation is made more explicit by the use of the relational predicate 'represents', along with the dual role given to atomic sentences, to both show and state something. On a coherence theory there is no correlation of atomic sentences to something extra-linguistic. This points to a basic defect of coherence theories.

On a correspondence theory, the condition for [aRb] being true is the existence of a fact. On a coherence theory, the truth condition is said to be the coherence of [aRb] with C. But, this raises a question as to what a truth condition is on a coherence theory. The expression '*coheres with*' appears to be a relational predicate, but the coherence theorist cannot take '[aRb] coheres with C' to represent *the fact* that the relation *coheres with* holds between [aRb] and C without abandoning the theory. Thus, such a theorist is forced to hold that the ground of truth for the proposition [aRb] is the truth of *the further proposition* that [aRb] coheres with C, and not the existence of a fact. But, according to the coherence theory, such a claim must be taken to mean that the truth condition for [aRb] is given by the statement:

[[aRb] *coheres with* C] *coheres with* C*,

where C and C* are possibly the same. Whether or not they are the same, a vicious regress arises, since the truth ground for a proposition (sentence) is always the truth of another proposition.[8]

The essential feature of a correspondence theory is its recognition of truth conditions that are not propositions or sets of propositions, since an existent fact provides a ground of truth. To object that the correspondence theorist assumes, in a question begging manner, that a ground of truth can only be a fact, and, hence, not a further proposition or set of such, is not cogent. All that is claimed is that though a coherence theory takes the ground of truth of [aRb] to be the coherence of [aRb] with C, such a theory cannot take the coherence of [aRb] with C to be an *ultimate* ground of truth. For, to do so is to accept a fact, rather than coherence with other propositions, as the ground of truth. The proposition [[aRb] coheres with C] cannot be taken to be true on the condition that there is a fact, that [aRb] coheres with C. Rather, it must be understood to be true on the condition that the proposition [[aRb] coheres with C], in turn, coheres with some set of propositions, C*, and so on ad infinitum.

The setting out of the correspondence theory in such a way as to show its conformity to Convention-T points to a misleading feature of Tarski's

linking of Convention-T, and his semantic conception of truth, with a correspondence theory. For, Tarski was not at all concerned either with the representational role of atomic sentences or with the analysis of their grounds of truth. A correspondence theorist, on the other hand, cannot be content to satisfy Convention-T, for such a theorist must be concerned with both the representational role of propositions and with facts as grounds of truth.[9] It is helpful to recall that a major concern regarding the introduction of a truth predicate into formal schemata, which is the context of Tarski's introduction of Convention-T, stems from Russell's attempt to resolve the liar paradox by means of his theory of propositional types. Subsequently, Russell suggested a "levels of language" approach in his introduction to the *Tractatus*. Tarski, as well as Carnap, developed such notions in a systematic way leading to important developments in formal semantics and logic.[10] But, traditional theories of truth, such as correspondence and coherence accounts, are another matter. These have to do with how one understands what truth conditions are, and not merely with questions about the introduction of a truth predicate, without paradox, into a formal language, where one simply deals with assignments of truth values to the basic sentences, i.e. with evaluations.

To give a theory of truth is to state what one takes a truth condition to be: it is not to give truth conditions in the sense that the truth condition for 'snow is white' is purportedly given by producing a token of the sentence 'snow is white' or of a metalinguistic transcription of that sentence. If a coherence theorist takes a truth condition to be the obtaining of a coherence relation, one can argue that the notion of a coherence relation obtaining is problematic. For a coherence theorist cannot take the obtaining of a relation to be a fact. To do so is to abandon the coherence theory. The coherence theorist must specify what a truth condition is in a way that is compatible with the coherence theory. The only alternative is to take the truth condition for a proposition, stating that a coherence relation obtains, to be a further true proposition, one in which the first proposition is itself the term of a purported coherence relation. This initiates the regress. It is reminiscent of the regress Russell once claimed that nominalist's face when they attempt to hold that a similarity relation is a particular and not a universal.[11]

As presented above, the argument against a coherence theory is also reminiscent of a criticism Frege directed against all theories of truth, and especially the correspondence theory:

> Can it not be laid down that truth exists when there is correspondence in a certain respect? But in which? For what would we then have to do to decide whether something were true? We should have to inquire whether it were true that an idea and a reality, perhaps, corresponded in the laid-down respect. And then we should be confronted by a question of the same kind and the game could begin again. So the attempt to explain truth as correspondence collapses. And every other attempt to define truth collapses too.[12]

Frege's argument can be taken to involve several different themes. One is that the correspondence theory, or any other theory purporting to analyze *truth*, faces a regress like that we directed against a coherence theory. The claim is that the correspondence theorist must offer '[[aRb] corresponds to aRb] is true' as the analysis of '[aRb] is true', and this initiates a regress, instead of providing a viable analysis. But, a carefully formulated correspondence theory does not do what Frege took it to do.

Consider (D3) as illustrating the way in which a correspondence theory employs a truth predicate in the context of presenting the theory. The atomic sentence 'aRb' is taken to refer to a situation, which provides one sense of the term 'correspond'. But, that it represents the situation it does is not a matter of fact. The sentence '[aRb] *represents* aRb' is true in virtue of the semantical rules of the schema for the bracketing device and the predicate 'represents'. It belongs to the logic of the schema. Hence, it is neither taken to be true in virtue of corresponding to a fact nor taken to *represent* a state of affairs.

There is another point to be made about Frege's challenge to a correspondence theory. As we have noted, the atomic sentence 'aRb' is used in two ways in (D3). In the first conjunct, it occurs as a term of the predicate 'represents', and stands for a situation irrespective of questions of truth or falsity. In the second conjunct, it functions as a sentence and not as a term of a relation. It is used there to assert that the situation it represents exists or obtains, and, hence, the obtaining of the situation is the ground of its truth. The right side of the biconditional, interpreted the way it is, anchors the atomic sentence to a non-linguistic item and thereby blocks one way in which such a regress might begin. We shall shortly consider the question of compound propositions on a correspondence theory of truth. If we assume, for the moment, that we give the truth condition for a conjunction by giving the truth conditions for its conjuncts, then, in the present case, the left conjunct of (D3) would be treated as an analytic truth. Hence, no question arises about a correlated situation or fact. The right conjunct is true if the situation represented obtains. If one takes conjunctions to represent conjunctive situations,

which could obtain or not, then a kind of regress of truth conditions could begin. It is not clear that it would be a vicious regress, as in the case of the coherence theory, or simply, like the Fregean generation of infinitely many senses from the sense of a single proper name, a case of having to acknowledge a further entity at any given stage. For, while it would be awkward to be forced to recognize an infinite number of facts, connected to each atomic sentence, one could hold that the existence of each member of such a sequence is entailed by the obtaining of the first situation involved, that represented by [aRb]. This circumstance does not arise for the sequence of truth conditions generated by a coherence theory. In the case of the coherence theory one does not merely conclude that there must be a further true proposition. Rather, since the coherence theorist must always appeal to a true proposition, as the ground of truth, for another, purportedly true proposition, such a theorist always fails to answer the original question. For the answer simply gives rise to exactly the same kind of question – a question about what constitutes the truth of a proposition. The point can be put in another way.

The correspondence theorist's explication of truth involves taking the claim that a subject *has* an attribute to employ two fundamental concepts – the connection combining the subject and attribute into a situation and the taking of such a situation or possibility to obtain or be actual. In effect, Frege argues that so putting matters involves taking such a claim to be true, which requires specifying a further truth condition. But, the correspondence theorist can safely hold that a fact's existing or obtaining is not itself a further situation that may or may not obtain, and so on ad infinitum. Though the coherence theorist is forced to accept an unending regress of coherence relations, there is nothing in the correspondence account of truth that forces one to hold that there is an unending series of situations or facts, generated by the atomic fact aRb. To put it another way, the fact that aRb (the obtaining of the situation that aRb) is the truth ground for each member of the series – 'aRb', "aRb' is true', "'aRb' is true' is true'.... , as well as of the series, if such there be, – it is the case that aRb, it is the case that it is the case that aRb, ...

Complications would arise if the correspondence theorist introduced a predicate transcribed by 'obtains'. Such a predicate would invite problems similar to those Bradley raised about a relation of exemplification. To see the connection, consider recent suggestions to nominalize predicates and employ a first order schema with a basic predication "predicate". One takes a predicate, 'G' for example, and instead of using

it as a predicate in predicate place, as in 'Ga', with 'a' as a name of a particular, one writes 'P(G, a)', with 'P' as a predicate expressing exemplification and 'G' and 'a' as subject terms. (It makes no difference, for this discussion, if, instead of replacing 'Ga' by 'P(G, a)', one introduces 'P(G, a)' as axiomatically equivalent to 'Ga' or uses 'P(g, a)', with 'g' as a "nominalized" form of 'G'.) The pattern is borrowed from the use of a first-order logic with a primitive membership relation in setting out an axiomatic set theory and the use of a many-sorted first order logic as an alternative to a second order logic in certain contexts. But, in contexts relevant to our concerns, such a pattern is open to the obvious objection that the syntactical arrangement of the signs in 'Ga' is used to represent the exemplification of a property by a particular, and it continues to be used in this way, if only covertly, in 'P(G, a)'. For, one takes 'P(G, a)' to state that a and G exemplify the relation P. The introduction of 'P' is thus both problematic and pointless for philosophical purposes. Similarly, to introduce a predicate 'obtains' and rewrite '[aRb] represents aRb and aRb' as '[aRb] represents aRb and aRb obtains' is pointless. It would still require one to hold that 'aRb obtains' does not itself represent a situation that may or may not obtain. Just as one can avoid Bradley's regress concerning predication, by holding that predication is a logical form and not a relation among relations representable by a predicate, one can avoid a similiar problem here by holding that a predicative connection is not involved and that 'obtains' is not a predicate among predicates. We can simply follow Wittgenstein and take atomic sentences, by themselves, to represent situations and to assert that situations obtain. For, even if one introduces a predicate like 'obtains', one still uses atomic sentences in two fundamentally different ways: in contexts with 'obtains' and in contexts without such a predicate. This brings to mind Wittgenstein's remarks about the uselessness of an assertion sign. Be that as it may, the correspondence theorist is forced to recognize a dual role for atomic sentences.

The link between Convention-T and the present sketch of a correspondence theory is based on the recognition of facts as truth grounds. We noted above that this gives rise to a major difference between a correspondence theorist's use of Convention-T and Tarski's use of that convention. The difference is vividly illustrated by something Tarski writes:

> If the language investigated only contained a finite number of sentences fixed from the beginning, and if we could enumerate all these sentences, then the problem of the

construction of a correct definition of truth would present no difficulties. For this purpose it would suffice to complete the following scheme: x∊Tr if and only if either x = x1 and p1, or x = x2 and p2,... or x = xn and pn, the symbols 'x1', 'x2', ..., 'xn' being replaced by structural descriptive names of all the sentences of the language investigated and 'p1', 'p2', ..., 'pn' by the corresponding translation of these sentences into the metalanguage.[13]

Tarski's understanding of Convention-T allows such a definition of truth to furnish an adequate analysis or theory without specifying what a truth condition *is* – a fact, a coherence relationship, a basic property of propositions, the denoting of a truth value, etc. In short, no analysis is given in the traditional sense. What is provided by Tarski's semantic definition, in terms of satisfaction, is an *extensionally equivalent* condition for applying a predicate to sentences of a schema that are evaluated in a certain way.

The contrast between a correspondence theorist's approach to Convention-T and Tarski's view can be vividly illustrated by considering an artificially simple case. Let us assume that we have only two atomic sentences, 'aRb' and 'bRa', in L, and no logical signs to enable the formation of non-atomic sentences. We can, then, construing single quotes in a standard way, and understanding the use of variables in a suitable way in the metalanguage, follow Tarski' suggestion and define a truth predicate by:

> x ∊Tr iff ((x = 'aRb' and a stands in R to b) or (x = 'bRa' and b stands in R to a)).

Thus, we get:

> 'aRb' ∊Tr iff (('aRb' = 'aRb' and a stands in R to b) or ('aRb' = 'bRa' and b stands in R to a)).

Since the identity in the second disjunct is trivially false, we obtain:

> (T') 'aRb' ∊ Tr iff ('aRb' = 'aRb' and a stands in R to b).

The conjunct "'aRb' = 'aRb'" plays the role of '[aRb] represents aRb', in the use of (D3) above. Just as '[aRb] represents aRb' was taken to be a theorem of the correspondence theorist's schema, and, hence, a logical or formal truth, the identity may be so taken.[14] Thus, as in the case of our correspondence theory, we obtain the requisite T-sentence and Convention-T is satisfied. The difference is that nothing at all is said about what it is that grounds the truth of 'aRb'. That is, nothing at all is said about the analysis of 'a stands in R to b', as philosophers have

traditionally sought such an analysis. This contrasts sharply with the analysis set out in a correspondence theory. There, by taking 'aRb' to represent a situation, the existence of which constitutes a ground of truth, one recognizes both a representational relation and facts, as well as the elements and structure of such facts.

The occurrence of 'aRb' and '[aRb]' in (D3), on the one hand, and the occurrence of "'aRb'" in (T'), on the other, serve to emphasize the contrast between Tarski's semantic theory and a correspondence theory. By providing a context of interpretation for (D3) that *specifies what a truth condition is* and *connects* 'aRb' to such a truth condition, the correspondence theorist offers an ontological analysis or assay. Tarski is not concerned to say anything further about 'a stands in R to b'. In fact, he seems to have thought that nothing intelligible remains to be said:

> I have heard it remarked that the formal definition of truth has nothing to do with 'the philosophical problem of truth'. However, nobody has ever pointed out to me in an intelligible way just what this problem is.... In general, I do not believe that there is such a thing as 'the philosophical problem of truth'.[15]

However, for a correspondence theorist, it is precisely the use of sentences like 'a stands in R to b' in (T') that raises the philosophical question about the ground of truth. Thus, as Aristotle may be taken to have done so long ago, a correspondence theorist purports to tell us what it is that grounds the truth of a sentence. Tarski showed that his type of definition of a truth predicate satisfied Convention-T, but such a definition does not capture the import of the Aristotelian conception of truth. A correspondence theorist proposes a theory of truth in another, traditional sense, that does give philosophical substance to the Aristotelian claim. In addition, such a theory, I have argued, not only satisfies Convention-T, while the coherence theory does not, but it provides a basis for a satisfactory ontological analysis by taking facts to be truth grounds. Tarski did not propose a philosophical or ontological analysis of the concept of truth or of the role of the predicate 'is true', as applied to sentences. But, he did set forth a condition that is philosophically significant. For, by setting out a criterion for satisfactory truth predicates, Tarski provided a partial explication of what it is to be a *truth predicate*, even if he did not provide an explication of what it is to be a *truth condition* or a *ground of truth*.

The difference between having a satisfactory truth predicate, in Tarski's sense, and having a satisfactory theory of truth can be further emphasized by modifying the simple case we considered above. Let us

take ‘a stands in R to b’, but not ‘b stands in R to a’, to be true, to be an axiom of our schema. We thus list the true atomic sentences of our schema, the only one being ‘aRb’. We then satisfy Convention-T by defining a truth predicate for L:

$$x \epsilon \mathrm{Tr} \quad \text{iff} \quad x = \text{‘aRb’}.$$

This would constitute a satisfactory definition of a truth predicate or theory of truth for L, in Tarski’s sense. Since ‘aRb’ is taken to be an axiom of the schema, we obtain the requisite T-sentence, for we have ‘aRb’ and:

$$\text{‘aRb’} \epsilon \mathrm{Tr} \text{ iff } \text{‘aRb’} = \text{‘aRb’},$$

where the right side is trivially (or logically) true. This means that in a sense we can be said to satisfy Convention-T, and hence give a satisfactory theory of truth for the schema, by listing the true sentences. Philosophical analyses are traditionally tested by considering extreme cases. Our simple case is such an extreme test case. A correspondence theory of truth does not turn out to be trivial in such a case, for it does more than offer a list as a theory. It claims that *there is something* that is a ground of truth for ‘aRb’, a fact. Likewise, our simple schema illustrates the fact that though Tarski’s Convention-T may be acceptable as a condition for a viable philosophical theory of truth or for the introduction of a truth predicate for other purposes, it is not acceptable as a sufficient condition for an analysis of the kind philosophers have traditionally sought, for it does not require that one specify what constitute the truth grounds for propositions. As we have noted, Tarski was not concerned with such matters. Rather, he was concerned with a formal question:

> We indicate which objects satisfy the simplest sentential functions; and then we state the conditions under which given objects satisfy a compound function–assuming that we know which objects satisfy the simpler functions from which the compound one has been constructed.[16]

Thus, to confine the point to the special case of sentences, Tarski’s concern was focused on the evaluation of complex statements, given the evaluation for the simplest sentences. This matter is relevant to the construction of a correspondence theory, for the correspondence theorist faces questions about the truth and representational role of non-atomic sentences. In fact, such a theorist may attempt to argue that one need not recognize non-atomic facts in virtue of the procedures Tarski employs

that, in a sense, reduce questions about the satisfaction of compound functions to questions about the satisfaction of simpler functions. Whether such a theorist can viably appeal to such procedures to avoid ontological commitments in connection with non-atomic propositions is a matter that we have not yet considered. So far we have only considered the correspondence theorist's fundamental claim that one who offers a theory of truth must face questions about the representational role of atomic sentences (propositions). Whether a correspondence theory need recognize non-atomic situations and facts is a further question that gives rise to alternative versions of such a theory.

We can get at this question by adding '$\vee$', with its standard interpretation, to the simple schema L. A correspondence theorist may then distinguish between the sense in which a disjunction is true and the sense in which an atomic sentence is true, by taking the latter, but not the former, to represent a situation. This difference could be expressed by the correspondence theorist holding, along familiar lines, that

(A) Tr[p $\vee$ q] iff (Tr[p] or Tr[q])

is the appropriate pattern specifying the truth (satisfaction) condition for a disjunction, with 'p' and 'q' as propositional variables and with the rules for brackets suitably understood to allow the use of variables. Thus, a disjunctive proposition is neither taken to represent a disjunctive situation nor to have a disjunctive fact as a truth ground. A correspondence theorist may thus use the different clauses in the standard recursive characterization of truth (satisfaction) for truth functional compounds like disjunction and conjunction to avoid recognizing correlated molecular situations and facts. Negation, however, is another matter.

Before turning to the problem posed by negation, it is worth recalling that correspondence theorists have often focused their concern about non-atomic facts on negation and generality.[17] Russell's claims about general and negative facts in the logical atomism lectures and in "On Propositions: What They Are and How They Mean," Wittgenstein's discussions in The *Tractatus* and his *Notebooks*, and the dispute between Ramsey and Moore are obvious examples.[18] The perennial challenge of the analysis of negative and false judgments is an aspect of the philosophical problem of truth that is ignored by Tarski's semantic conception of truth, as is clear from Tarski's comment: "Therefore every definition of truth which is materially adequate would necessarily be equivalent to that actually constructed." This statement emphasizes the

limitations of the appeal to Convention-T as a criterion and of the semantic conception of truth as a theory, even when the issue is confined to a dispute about negative facts, between adherents of different versions of a correspondence theory of truth. For Tarski any predicate that is *coextensive* with the truth predicate he has defined (for the particular schema in question) will do. And this is merely a matter of *what* statements are taken to be true and not a matter of the nature of *the grounds of truth*.

In dealing with negation for atomic sentences three choices arise for a correspondence theory of truth. They are illustrated by the following biconditionals:

(a) Tr[¬ aRb] iff ([¬ aRb] represents not-aRb and not-aRb)

(b) Tr[¬ aRb] iff ([aRb] represents aRb and not-aRb)

(c) Tr[¬ aRb] iff ¬ Tr[aRb] iff it is not the case that ([aRb] represents aRb and aRb).

We can take the use of 'not' in the left conjunct in (a) to make explicit that to employ (a) is to take a negated atomic sentence to represent a *negative* situation, which then exists or does not. This would mean that negative situations, *as well as* non-existent situations, are acknowledged. Wittgenstein, inconsistently, seems to have implicitly expressed such a view at places in the *Tractatus*.[19] Such a view is forced to make a four-fold distinction among situations, deriving from the dichotomies positive–negative and possible (potential)–actual (existent). But, there is no need for a correspondence theorist to so extend his ontology. Thus, (b) embodies the view that to say that [¬ aRb] is true is simply to say that the situation respresented by [aRb] does not exist. (b) employs the English (meta-linguistic, if one prefers) term 'not' in that sense, just as 'aRb', in the right hand conjunct of '[aRb] represents aRb and aRb', is used to state that the situation *does exist*. This is what (c) amounts to as well. For, as '[aRb] represents aRb' is a linguistic or logical truth, '¬ Tr[aRb]' reduces to 'not-aRb', used, as in (b), to state that aRb does not exist (that a does not stand in R to b). In this way the construal of situations enables the correspondence theory we are developing to deal with the problem of negative facts. All we need recognize are situations that exist and situations that do not exist, and these are already implicitly recognized in specifying the truth conditions for atomic propositions.

The role of negation in specifying truth conditions points to a way in which negation is unlike other truth functions of atomic propositions. For, while atomic propositions are taken to represent situations and to assert that they exist (are facts), a negation of an atomic proposition is taken to assert that the situation, represented by the constituent atomic proposition, does not exist. Thus, the standard scope reduction of a negation, employed in dealing with negations of non-atomic statements, reflects the dual role negation plays, since negated molecular propositions "reduce" to cases of negated atomic propositions or cancel out, but negated atomic propositions have non-obtaining situations as truth grounds.

The distinction between speaking of the truth of atomic sentences, where we take facts as truth conditions, and the ascription of truth to molecular statements, where molecular statements are not held to correspond to complex facts, bears on recent concerns about the ambiguity of a truth predicate and the liar paradoxes. In the years between 1906 and 1910, Russell set out a theory (various theories, actually) of types for propositions to avoid self-referential paradoxes and "vicious circle fallacies," such as "the liar" or "the Epimenides." In the introduction he wrote for The *Tractatus*, he suggested a variant of this approach for sentences, involving a hierarchy of languages, with an appropriate truth predicate for each level.[20] Such an approach obviously does not fit with the fact that ordinary languages are not typed in such a manner. Thus, some seek to avoid the familiar puzzles and paradoxes without introducing a hierarchy of truth predicates, and the consequent "systematic" or "typical" ambiguity of truth.[21] The basic idea behind such attempts stems from Frege's introduction of truth value gaps to resolve familiar puzzles about non-denoting singular terms, Lukasiewicz's later use of the same device in connection with statements about the future in order to avoid "determinism," and Strawson's subsequent reiteration of such a theme in his criticism of Russell's theory of definite descriptions.[22] Thus, in various ways that give rise to different so-called "theories of truth," one avoids the liar paradox, and its variants, by taking the crucial sentences either to be without a truth value, or to shift truth values on different "evaluations," or to obtain a truth value relative to an evaluation, etc.

While such alternatives do not employ a hierarchy of truth predicates, they do employ alternative devices of gaps, shifts that change the extension of a truth predicate, etc. devices that neither Russell nor Tarski

had to introduce. Thus, one issue that arises concerns the technical effects of the various restrictions imposed by typed as opposed to non-typed theories. Another issue concerns the purported affinity with natural languages possessed by some, but not other analyses. What is important for a viable ontological analysis, over and above formal concerns about the introduction of a truth predicate into a schema, is that Russell's theory of propositional types, and the sentential variant involving language levels that he later suggested, reflect a fundamental theme of a viable correspondence theory of truth. Atomic sentences are true in virtue of representing situations that obtain. Molecular sentences are true under different sorts of truth conditions, though their truth "reduces" to that of atomic statements.

Russell advocated a theory of types for truth concepts to reflect this difference, as well as to avoid familiar paradoxes. For Russell, a truth predicate ascribed to molecular propositions was of a different type than that ascribed to atomic propositions, since the former reduced to a statement involving ascriptions of truth to atomic propositions, while the ascriptions to atomic propositions reduced, in turn, to assertions that corresponding complexes existed. Whether one holds the standard view that we have one truth predicate, with various clauses in its recursive definition (or in the definition of 'satisfaction' in terms of which the truth predicate is defined), or prefers to follow Russell in using types since truth claims about molecular propositions reduce to truth claims about constituent propositions, while truth claims about atomic propositions are understood in terms of claims *about the existence and non-existence of situations*, and not about other propositions, matters little. What does matter is the recognition of the link between atomic propositions and their negations, on the one hand, and the existence and non-existence of situations, on the other.

Given the general association between theories of types and the familiar paradoxes, it is worth recalling how Russell and Whitehead introduced the type distinction, for propositions, in *Principia Mathematica*:

> We will give the name of "a complex" to any such object as "a in the relation R to b" or "a having the quality q," or "a and b and c standing in the relation S." Broadly speaking a complex is anything which occurs in the universe and is not simple. We will call a judgment elementary when it merely asserts such things as "a has the relation R to b," "a has the quality q" or "a and b and c stand in the relation S." Then an elementary judgment is true when there is a corresponding complex, and false when there is no corresponding complex.
>
> But take now such a proposition as "all men are mortal." Here the judgment does not

correspond to one complex, but to many, namely "Socrates is mortal," "Plato is mortal," "Aristotle is mortal," etc....

It is evident ...that the definition of truth is different in the case of general judgments from what it was in the case of elementary judgments. Let us call the meaning of truth which we gave for elementary judgments "elementary truth." Then when we assert that it is true that all men are mortal, we shall mean that all judgments of the form "x is mortal," where x is a man, have elementary truth. We may define this as "truth of the second order" or "second-order truth."[23]

Thus, Russell and Whitehead took the notion of different truth concepts, or different definitions of truth, to stem from the different accounts of truth conditions for "elementary" (here, atomic) propositions as opposed to non-atomic propositions, and not simply as a means of avoiding the semantic paradoxes. Only atomic propositions were taken to correspond to "one" complex or fact, that is, to *a* complex or fact. This claim fit with Russell's continuous concern about negative, compound, and general facts, which he later attempted to resolve, in the logical atomism lectures of 1918, by acknowledging both negative facts and general facts, while rejecting facts corresponding to molecular statements.[24] In *An Inquiry into Meaning and Truth* he took molecular statements to belong to a metalanguage, for an "object" language containing atomic statements.

While the semantic paradoxes provided another reason for his considering the concept of truth to be *systematically ambiguous*, Russell's version of a correspondence theory of truth provided an independent basis for recognizing such a systematic ambiguity. Recent attempts to reject types and avoid the paradoxes by employing a variant of a truth value gap analysis, that purportedly allows for a univocal truth predicate, illustrate problems for such an approach for philosophers concerned about ontological issues and truth grounds, and not merely about formally adequate truth predicates. The acceptance of truth value gaps introduces a purportedly univocal truth predicate at the price of introducing an alternative systematic ambiguity, an ambiguous construal of predication, which is problematic for a correspondence theory.

On the version of the correspondence theory of truth we have considered, an atomic subject-predicate statement is taken to represent a situation and be true if that situation obtains (is a fact). This *presupposes* that the predicative juxtaposition of subject and predicate terms is taken to represent the exemplification of a property by a particular, the former represented by the predicate and the latter represented by the subject term. To allow for truth value gaps for atomic sentences is to abandon

such an explication of linguistic predication. Moreover, the acknowledgment of propositional entities, whether explicitly taken to be such or taken as Fregean thoughts, Moore's beliefs, or Strawson's assertions, is of no help. For, aside from introducing problematic entities, the original problem persists, since truth conditions must now be specified for such propositional entities, rather than for statements. Here, I am assuming, and can not argue, that Frege's introduction of *The True* and *The False* merely avoids the problem of specifying truth conditions, in much the way Tarski's semantic conception of truth does.[25]

The point is that the predicative tie in a propositional entity, which is expressed by the predicative connection of terms in a sentence, is understood to represent an exemplification relation in a purported fact or situation. To introduce truth value gaps is thus to abandon a univocal understanding of predication. This is indeed ironic if the motive for introducing truth value gaps is to have a univocal truth predicate. Yet, while it might be granted that such an argument applies to attempts, by Strawson and others, to take atomic sentence patterns to have truth value gaps, one might object that no atomic sentence need be taken to have a truth value gap in order to deal with liar type paradoxes. And, since sentences employing a truth predicate are not taken to represent situations or facts, on a viable correspondence theory, there is no need to recognize the ambiguity of predication. For, given that such statements do not represent situations, no question arises about the ambiguity of predication. This objection points to a crucial question about truth predicates.

The correspondence theory that we have set out neither takes a truth predicate to represent a property nor takes a sentence like 'Tr[aRb]' to represent a situation that contains [aRb], whether the latter is construed as a sentence or a propositional entity. Thus, the use of such a predicate does not involve the ascription of a property in the way that the use of a primitive predicate in an atomic statement does. This is simply a way of putting the point that in the case of an atomic proposition or statement [aRb], the truth condition for 'Tr[aRb]' is simply the truth condition for [aRb]. In order to specify the truth condition for a statement or proposition employing a truth predicate (concept), one must specify the truth condition for a statement or proposition without that predicate or concept. This point is captured by a theory of types in that the lowest type of statement cannot contain a truth predicate. It is reflected in supposed type free accounts, such as Saul Kripke's, by having the evaluations of

sentences containing the truth predicate require prior evaluations of other sentences, and if the required prior evaluations do not occur, such sentences are not evaluated. Liar sentences then become sentences that retain truth value gaps through successive evaluations that extend the extension of the truth predicate.

However, on such non-typed gap theories that allow liar sentences, sentences involving a truth predicate need not "reduce" to ascriptions of basic predicates. This is why such theories either do not evaluate liar sentences or have "levels" of assignments shifting the value. This aside, in that statements with a truth predicate need not "reduce" we have an ambiguous construal of predication, just as on accounts, like Strawson's, that allow for truth value gaps in the case of atomic sentences with non-naming names. To accept non-naming names and truth value gaps for atomic sentences is to abandon the explication of predicative combination whereby one takes a monadic subject-predicate sentence to state that the object represented by the subject sign, in fact, exemplifies the attribute represented by the predicate and a relational sentence to state that the objects represented by the subject signs, in fact, exemplify the relation represented by the relational predicate, in the represented order. That is, one who employs truth value gaps, as Strawson does, to resolve puzzles about reference and denotation abandons such a construal of predication in favor of an equivocal conception of predication. For sentential predication amounts to one thing with non-denoting subject terms and another thing with denoting signs. Kripke's use of truth value gaps, to purportedly obtain a single, univocal truth predicate, does so by employing an ambiguous construal of predication in a similiar manner.

The point is emphasized by Kripke's claim that a modified version of Convention-T is satisfied by his variant of a non-typed schema.

> The approach adopted here has presupposed the following version of Tarski's "Convention-T", adapted to the three-valued approach: If 'k' abbreviates a name of the sentence A, T(k) is to be true, or false, respectively iff A is true, or false. This captures the intuition that T(k) is to have the same truth conditions as A itself; it follows that T(k) suffers a truth-value gap if A does.[26]

This highlights the fact that an atomic sentence can be of the form 'T(k)', in Kripke's schema, while on a Russellian style type theory, a truth predicate cannot occur in a sentence that is atomic in form and irreducible. Moreover, given our correspondence theory's pattern for the specification of truth conditions for atomic sentences, it is clear that for

a correspondence theory to treat a truth predicate as Kripke does would require that the predicate be taken to represent a property. This provides a sharp contrast with a Russellian type approach, for one who seeks to provide a theory of truth in a classical sense, along with an ontological assay of truth grounds, and is not merely concerned with the introduction of a truth predicate that avoids paradoxical contexts. It also reveals a two-fold problem. For, not only would the construal of 'T(k)' in terms of:

(Kr) 'T(k)' represents T(k) and k has T,

force one to take 'T' to represent a basic property, it would also require one to deny that 'k has T' represents a situation, which either exists, and thus furnishes the truth ground for the atomic sentence 'T(k)', or does not. (That one could take it to represent a situation which was not "bivalent" would not alter the point.) This reflects an ambiguous use of 'has': its use in clauses specifying the truth conditions for "normal" atomic statements and its use in sentences like (Kr). If one does not treat 'T(k)' in terms of (Kr), then one explicitly acknowledges a distinction among atomic sentences regarding the role of predicative juxtaposition. In either case, we have an ambiguous construal of predication in atomic sentences. One must keep in mind the different ways the notion of a "truth condition" is used–as Kripke uses it in the above quotation and as it as construed by a correspondence theory. The simple point is that if one allows the kind of sentences Kripke allows, irrespective of a construal of 'is true' in terms of a satisfaction relation, the relevant truth condition is expressed by 'k has T'.

Kripke's suggested modification of his schema, which amounts to taking 'T(x)' to be false of any sentence that has not been evaluated as true, at a certain point of the evaluation process, does not alter this point.

> A modified version of Tarski's Convention T holds in the sense of the conditional $T(k) \vee T(neg(k)).((. \supset .)) A \equiv T(k)$. In particular, if A is a paradoxical sentence, we can now assert $\neg T(k)$. Equivalently if A has a certain truth value before T(x) was closed off, then $A \equiv T(x)$ is true.[27]

This does not change matters since we are concerned with the problem of specifying what entities provide the truth grounds for atomic sentences, and atomic sentences containing 'T' are construed differently from normal atomic sentences. Moreover, as Kripke notes, he is now forced to recognize a metalinguistic truth predicate differing from his "object

language" truth predicate.[28] Once one does that the philosophical significance of the analysis is called into question on familiar grounds, grounds that led Russell to reject purported nominalistic accounts of predication that apparently require *only one universal* similarity relation.

Russell and Whitehead, as we noted, took quantified statements to require a different truth predicate than that ascribed to atomic sentences. Such statements, which led Russell to speak of general facts, must be dealt with on a correspondence theory recognizing situations or possibilities. Whether one can simply make use of the familiar characterization of a satisfaction relation without recognizing additional kinds of facts and situations, general and complex situations and facts, is an obvious question that arises. If not, one would be obliged to provide an analysis of such entities, which might require the recognition of logical forms, represented by '$(x)\hat{\Phi}\, x$' and '$(\exists\, x)\hat{\Phi}\, x$', that combined with properties to form situations and facts. But, however one handles such complex contexts, two fundamental problems remain for such a variant of the correspondence theory.

Consider the statement:

[Fa] represents Fa and Fa.

Given our variables, we can existentially generalize to obtain:

($\exists$ p)([Fa] represents p and p).

From this, we can proceed to:

($\exists$ p)p.

Earlier, we read this statement as 'Some situation obtains'. One might find such a statement odd, even though we read 'Fa', in the right conjunct above, as the claim that a situation obtains (exists, is positive). The objection is really an objection about the two-fold use of atomic sentences: to represent situations and to assert that they obtain. Consequently, we use the variable 'p' in clauses like '[Fa] represents p and p' to range over situations and to assert that they obtain (or do not).

The type of correspondence theory we have developed, derived from Moore, Wittgenstein, and Russell (of 1918-20), makes use of an ambiguous sense of 'represents'. Primitive constants, like 'a' and 'F', are coordinated to an object and property respectively. (Russell, recall, thought that since the entitites were so different, the connection between a name and a particular was different from that between a predicate and

a property, while the connection between an atomic sentence and a complex was different yet.) Atomic sentences are not simply coordinated to situations, for while the names and predicates may be taken as labels of their correlated objects, a sentence is hardly a label. This was the point of Russell's well-known comment that Wittgenstein had taught him that sentences were not names. The double use of atomic sentences in conjuncts like '[Fa] represents Fa' and 'Fa' reflects this two-fold use of 'represents' and poses a further problem. Such problems, along with the recognition of situations, either as constituents of positive and negative facts or as obtaining and non-obtaining, suggest an alternative variant of a correspondence theory–a variant Russell attempted to develop over the years 1905-1913 and abandoned in the period of the logical atomism lectures.

One can avoid possibilities by specifying truth conditions for atomic sentences in terms of existential claims, with quantifiers ranging over *existent* atomic facts. Retaining the relation P and sentential variables, but taking the former as a relation between an atomic *fact* (hence an existent) and a constituent of it, while the latter are restricted to the set of (existent) atomic facts, we can specify the truth condition for 'Fa' as follows:

$$\mathrm{Tr}[\mathrm{Fa}] \text{ iff } (\exists\, p)(P(F, p) \text{ and } P(a, p)).$$

In the case of negation we can use:

$$\mathrm{T}[\neg\, \mathrm{Fa}] \text{ iff not-}(\exists\, p)(P(F, p) \text{ and } P(a, p)).$$

The idea is that by recognizing the set of atomic facts, we can specify a truth condition in terms of the assertion, or denial, that there is a fact with certain constituents. Such a view enables a correspondence theorist to clearly separate the sense in which a sentence represents something and the sense in which a constituent primitive term represents a particular or a property. In the latter cases we simply have a stipulated coordination and can take sentences like "'a' represents a" and "'F' represents F" to be formal truths that are instances of the analytic pattern "'...' represents ...", where the same name or primitive predicate replaces both occurrences of the dots. By contrast, to speak of an atomic sentence representing a fact is to speak of the representational role of its consituent terms and the connection set forth in the definition of the truth predicate for atomic sentences. We thus avoid an equivocal use of 'represents'.

One must, as on the theory recognizing situations, specify the truth

conditions for quantified statements, but given that we have acknowledged domains of facts, of particulars, and of properties, such contexts can be handled in a straight forward manner. For primitive predicates, the following patterns are available:

$$\mathrm{Tr}[(\exists x)Fx] \text{ iff } (\exists x)(\exists p)(P(F, p) \text{ and } P(x, p))$$

$$\mathrm{Tr}[(x)Fx] \quad \text{ iff } (x)(\exists p)(P(F, p) \text{ and } P(x, p))$$

with obvious treatments of negation. One might add Russell's logical form of monadic predication, represented by '$\hat{\Phi}\hat{a}$' as a further constituent of such facts, to ward off certain objections and achieve kind of closure. One might even take '$(\exists x)Fx$' to be rendered in terms of '$(\exists p)P(F, p)$', given that 'p' ranges over the domain of atomic facts, which does not include higher order facts. Hence, F can only enter into an atomic fact in "one way" as the attribute of some particular also in the fact. But these details need not be taken up here. For quantification over complex contexts, such a theory can make use of the familiar recursive construal of satisfaction, as in the case of molecular compounds.

While such an alternative avoids the features of a correspondence theory that acknowledges situations and negative facts, the use of atomic sentences in a two-fold manner and, consequently, the use of an ambiguous construal of 'represents', it faces a problem of its own. For it would appear that it cannot satisfy Convention-T for atomic sentences, as the theory recognizing situations can satisfy that convention. Hence, in one way it shares a defect of coherence theories. The appearance is deceptive.

The problem is to derive a T-sentence without adding an assumption, in the case of 'Fa', like:

$$Fa \text{ iff } (\exists p)(P(F, p) \text{ and } P(a, p));$$

the sort of assumption that the correspondence theory recognizing situations does not require. But the present theory does not require such an assumption; it follows from the theory's account of predicative juxtaposition. As we noted in our earlier discussion of truth value gaps, a correspondence theory construes predicative juxtaposition, as in the case of 'Fa' for example, by understanding such an arrangement in terms of "There is an x and an f such that 'a' represents x and 'F' represents f and x has f". Thus, a correspondence theory not only introduces a truth predicate, it also specifies how predicative juxtaposition is construed.

These obviously have to be formally connected, but they are not literally the same thing. The link is provided by the way the theory construes the clause 'x has f'. On the correspondence theory that recognizes situations, this is done by the two-fold use of atomic sentences in expressions like "'Fa' represents Fa and Fa". On the present alternative, recognizing only atomic facts, 'x has f' is understood in terms of '($\exists$ p)(P(f, p) and P(x, p))'. Thus the construal of monadic predicative juxtaposition for atomic sentences is expressed by:

(CT) Fa iff ($\exists$ x)($\exists$ f)('a' represents x and 'F' represents f and ($\exists$ p) (P(f, p) and P(x, p))).

Since the right hand of the biconditional is an existential generalization from – 'a' represents a and 'F' represents F and ($\exists$ p)(P(F, p) and P(a, p)) – and since "'a' represents a", "'F' represents F", "($\exists$ *x*) ('a' represents x iff x = a)", etc. are formal truths of the schema, we easily obtain both 'Tr[Fa] —> Fa' and 'Fa —> Tr[Fa]' Thus, on the present version of a correspondence theory, atomic sentences satisfy Convention-T. We have, however, recognized sets or domains over which the variables 'x', 'f', and 'p' range: sets of objects, of primitive properties, and of atomic facts. The recognition of such sets allows us to specify the truth grounds for general statements, like '(x)Fx', without recognizing a Russellian "general fact" claiming either that the objects are all the objects or that the atomic facts are all the atomic facts. By recognizing the set of all atomic facts, we furnish the ontological ground for such general claims. This requires some comments about the ontological assay of sets or classes on such a correspondence account of truth.

On the present correspondence theory, we have a domain or set of particulars, one of primitive properties, and one of facts. These elements are like the "ur" elements of some set theorists. However, unlike set theorists, in that we are setting out an ontological assay of truth grounds, we are obliged to state how we construe these further entities, i. e. to state what sets are. Just as the logical form of monadic exemplification or $\hat{\Phi}\hat{x}$, is a connection or nexus, there is also a combinatorial connection represented by the braces '{ }' that connects particulars into complexes. It also can be taken to form further complexes out of such complexes, with obvious restrictions. The same holds for properties and atomic facts, that is for the various entities of the ontology. Such complexes are classes or sets. They are assumed to exist, irrespective of whether we describe or specify them or form signs for them, and they form something like a

cumulative hierarchy. The connection, which we may call the 'set operation', gives rise to complexes in much the way exemplification gives rise to facts (or situations, on the alternative theory), though there are fundamental differences. First, where we have entities like a and b or a and F, we have the classes {a, b} and {a, F}; we do not, by simply having entities that can be components of atomic facts, have such facts. Second, facts require certain kinds of components, while sets, apart from restrictions needed to avoid the paradoxes, do not. Third, the set operation can combine with, or operate on, one entity or argument, unlike the logical forms of monadic exemplification, dyadic exemplification, etc., which are all relational. Thus, the set operation need not connect many elements into "one". Fourth, the set operation operates on infinitely many terms or arguments, and, hence, cannot in principle be represented, in a normal schema, by a juxtaposition of signs representing terms.

The operation on, or combination with, one or more terms results in a class or set. This is taken to be composed of the term, or terms, and the set operation. If we have only one term, the result is a unit set. We can then consider the empty class or set to be the set operation itself, and, hence, the only class that is not a complex. This provides a way of *ontologically* grounding both the notion of an empty class and the difference between an element and its unit class. A fundamental mistake sometimes made by Russell was to think of a mere description of a set, by means of a property or a propositional function, as being essential to the existence of a set. Since there are objects, which, for example, have the property F, there is a certain set. But properties, or functions, and expressions like '{x| Fx}' simply enable us to describe a set. Neither the description nor the property has anything essentially to do with the existence of the set. That there are sets that can only be described by specifying a property of all elements is an interesting fact about sets, nothing more. That certain purported descriptions lead to perplexity and paradox is another interesting matter.

The very way we have characterized a set precludes anything from being identical with its own unit set, since the analysis of a unit set involves the set operation in a way that its element does not. In short, one makes use of the basic idea involved in hierarchies designed to dispense with the paradoxes. The point is that this is based on the way in which sets are construed. Of course, put as axiomatic theorists do, we are in effect specifying axioms for such a set operation. The fundamental point is that

given two particulars, we have something else, a set of them, given three, a set of them, and so on. What grounds the existence of such further entities is the combinatorial connection { }. This serves to resolve Russell's old perplexity about a class as "one" and a class as "many". We have a class as one in that every class, other than the null class, is a complex entity; we have a class as many where we have classes constituted by that connection combining more than one element. We can also see why a truth like 'a ϵ {a, b}' is trivial, while one like 'a ϵ {x | Fx}' need not be. The first is used to assert that a is an element in a set that is described as resulting from the set operation connecting a and b, while the second is used to state that a is an element of a set that is described without reference to any specific elements, but simply as the complex that results from the set operation combining the objects that have F. There is a fundamental difference between the exemplification nexus $\hat{\Phi}\hat{x}$ and the set operation { } that is worth noting. Exemplification, it was argued earlier, is a basic logical form or relation that is not viably representable by a predicate. Set membership is not a relation at all. The sign 'ϵ' is simply a linguistic reflection of the basic set operation. The truth condition for 'aϵ {a, b}' is not the obtaining of a relation between a and {a, b}. Rather, it is the existence of {a, b} that is the truth maker for the sentence. This is why the sentence appears trivial and redundant. By contrast, the truth of 'aϵ {x | Fx}' is grounded in two things: the fact that a is F and the formation of sets from particulars.

Given the atomic facts, as truth makers, we have the set of all atomic facts. This set is taken to exist without taking its members to have a property – being an atomic fact. That an atomic fact is such is no different from a particular being a particular or a property a property. One may say this is a matter of logical form and not a question about further matters of fact. This being so, there is no problem, on the alternative correspondence theory we have developed, either about using quantifiers like '(p)' over the domain of facts or of generating an infinite series of facts from any given atomic fact. To say that it is a fact that there is an atomic fact containing a and F is simply to say that there is an atomic fact containing a and F.*

NOTES

[1] A. Tarski, "The Concept of Truth in Formalized Languages," reprinted in A. Tarski, *Logic, Semantics, Metamathematics*, trans. J. H. Woodger, (Oxford: 1956), pp. 152-278.

On the dating see the bibliographical note, p. 152 and the footnotes on p. 154 and pp. 247-248.
[2] Tarski, 1956, pp. 187-188.
[3] Tarski, 1956, p. 155 and A. Tarski, "The Semantic Conception of Truth," *Philosophy and Phenomenological Research*, 4, 1944, pp. 342-343.
[4] Tarski, 1944, p. 343.
[5] Tarski, 1944, p. 356.
[6] L. Wittgenstein, *Tractatus Logico-Philosophicus*, trans. by D. F. Pears and B. F. McGuinness, (London: 1961), p. 41. We should note that an alternative way of construing a correspondence theory that appeals to situations is to take situations to be constituents of positive or negative facts. Thus, instead of recognizing existent and non-existent facts or situations, one recognizes positive and negative facts, both of which are existents, that contain constituent situations, much as Aristotelian substances, in one sense, contain prime matter. Just as prime matter is not an entity, but involved in the analysis of entities, so possibilities would not be entities but involved in the analysis of entities. This distinction, while of little importance for our purposes, can be important in discussions of purported possible worlds.
[7] I am quite aware that Wittgenstein struggled to avoid the "shadowy" entities we are taking as situations or possibilities, by taking the possibilities to be given by the natures or "essential" or "internal" properties of objects. On this see H. Hochberg, "Facts, Possibilities and Essences in the *Tractatus*," in E. D. Klemke ed., *Essays on Wittgenstein* (Urbana: 1971), pp. 530-533. The question is whether Wittgenstein's appeal to internal properties of objects and corresponding internal properties of their representatives, in thought or in language, carry the burden of correlating a complex, a proposition, to its truth condition. For, the correlation of the respective constituents will not do unless one takes it to be internal to being a sign, 'a', for example, or a token of a sign, not only that it can combine with another, 'F' say, in a certain arrangement, but that the resulting arrangement represents the situation that a is F. But this is not to correlate complexes by correlating their constituents, it is to correlate complexes directly, though it is *dependent* on the correlation of constituents. For, it amounts to correlating the arrangement 'Fa' to the situation. The internal property that the sign has is that it can combine with another in an arrangement that is understood to represent the situation. In any case, little is gained by giving objects such essential or logical properties in order to avoid possible facts. At best, one ends up with a having the internal property of *possibly being F*, instead of having a be a constituent of the situation represented by 'Fa'.
[8] That the coherence theorist can take [[aRb] coheres with C] and [[[aRb] coheres with C] coheres with C*] to have the same truth value is irrelevant. The point is that he must hold that the truth ground for one proposition is the truth of another proposition. The regress is blocked by taking $C = C^*$, holding that the propositions are logically equivalent, and that logically equivalent propositions are identical. But this amounts to making the proposition its own ground of truth.
[9] In spite of some passages in Tarski's original paper and the way the "semantic conception" of truth is sometimes discussed, as, for example, by E. W. Beth in *The Foundations of Mathematics*, (Amsterdam: 1959), p. 340:

"..the clause "t is in K" explicitly describes the state of affairs which the sentence T(t*) of A is to express. The statement (T) strongly resembles traditional definitions of truth as adaequatio rei et intellectus."

it is misleading to link Tarski's views with the correspondence theory. In Tarski, 1944, p. 343, Tarski quite explicitly holds that taking sentences to designate "states of affairs" is the kind of formulation that leads "to various misunderstandings" since it is not sufficiently "precise and clear".

[10]For Carnap's version of a linguistic hierarchy and Convention-T, see R. Carnap, *The Logical Syntax of Language*, (London: 1937), especially pp. 205-32.

[11] B. A. W. Russell, "On the Relations of Universals and Particulars," *Proceedings of the Aristotelian Society*, 1911-12, pp. 1-24.

[12] G. Frege, "The Thought: A Logical Inquiry," trans. by A. M. and M. Quinton, Mind, 65, 1956, reprinted in *Essays on Frege*, ed. E. D. Klemke, (Urbana: 1968), p. 510. Frege's argument is reminiscent of F. H. Bradley's attack on all relations, including a correspondence relation.

[13] Tarski, 1956, p. 188.

[14] In the present discussion I follow Tarski's use of 'ϵ'. It is worth noting that Tarski offers a definition of the "semantical concept" designates such that: ϕ designates π iff $\beta = \pi$, where 'ϕ' is a metalinguistic name of the sign 'β'. Tarski, 1944, p. 373, n. 2. If we apply this pattern to sentences like 'aRb', by taking the sentence to name something, a situation or state of affairs, while treating the occurrence of the sentence in quotes as a name of the sentence, we get: 'aRb' designates aRb iff aRb = aRb.

[15] Tarski, 1944, p. 361.

[16] Tarski, 1944, p. 353.

[17] On such questions see H. Hochberg, "Negation and Generality," *Nous*, 3 1969, pp. 325-43, and "Russell, Ramsey and Wittgenstein on Ramification and Quantification, *Erkenntnis*, 27, 1987, pp. 257-81.

[18] See G. E. Moore, "Facts and Propositions," *Aristotelian Society Supplementary*, vol. vii, 1927, reprinted in *Philosophical Papers*, (London: 1959), pp. 60-88; F. P. Ramsey, "Facts and Propositions," *Aristotelian Society Supplementary*, vol. vii, 1927, reprinted in F. P. Ramsey, *The Foundations of Mathematics*, ed. R. B. Braithwaite, (London: 1931), pp. 138-55, for the Ramsey-Moore debate.

[19] For example, 4. 0621.

[20]B. Russell, "Introduction," in Wittgenstein, 1961, p. xxii.

[21] See, for example, S. Kripke, "Outline of a Theory of Truth," *The Journal of Philosophy*, 72, 1975, pp. 690-716; H. G. Herzberger, "Notes on Naive Semantics," *Journal of Philosophical Logic*, 11, 1982, pp. 61-102; A. Gupta, "Truth and Paradox," *Journal of Philosophical Logic*, 11, 1982, pp. 1-60. All three papers are reprinted in R. L. Martin, *Recent Essays on Truth and the Liar Paradox* (Oxford: 1984).

[22] For Lukasiewicz's use of truth value gaps, see J. Lukasiewicz, "On Determinism," in *Polish Logic*, 1920-1939, ed. Storrs McCall, (Oxford: 1967), pp. 19-39. The ascription of the idea of truth value gaps to Frege is based on his discussion of fictional names in "On Sense and Reference," his classification of thoughts as true, false, or fictitious in an 1897 logic manuscript, and his stating that the sentence 'the sum of the Moon and the Moon is one' is neither true nor false in his *Grundgesetze*. However, he might merely have meant that non-denoting singular terms and sentences containing them ("transparently") need not be given serious consideration. I am indebted to Ignacio Angelelli for this cautionary note.

[23] A. N. Whitehead and B. A. W. Russell, *Principia Mathematica*, vol. 1, 2cnd. ed., pp. 44-45.

[24]B. A. W. Russell, "The Philosophy of Logical Atomism," reprinted in *Logic and Knowledge: Essays 1901-1950*, ed. R. Marsh, (London: 1956), pp. 236-37.
[25] One reason for Frege's introduction of the objects The True and The False, as basic objects denoted by propositions ("thoughts") and sentences, may well have been his view that no analysis or theory of truth was viable. Hence, one could only state that a proposition was true or false – denoted The True or The False.
[26] In Martin, 1984, p. 80. In the above discussion I have ignored the irrelevant complication that non-paradoxical self-referential statements can be given arbitrary evaluations, instead of retaining truth-value gaps.
[27] In Martin, 1984, pp. 80-81.
[28] Thus he writes "The necessity to ascend to a meta-language may be one of the weaknesses of the present theory." In Martin, 1984, p. 80.

* I have profited from discussions with Nicholas Asher and Per Lindstròm. This is not to say that either of them would agree with anything written above.

GRAEME FORBES

WORLDS AND STATES OF AFFAIRS: HOW SIMILAR CAN THEY BE?

1. MOTIVATION

A state of affairs is a partial world. Alternatively, a world is a maximal state of affairs. In this paper, I want to consider a proposal due to Nathan Salmon about how fine-grained the ontologies of worlds and states of affairs are.[1] One natural way of thinking of worlds and states of affairs is as constructs out of objects and properties.[2] I will argue that Salmon's proposal fails because it presupposes an incoherent conception of what an object is.

Salmon's proposal is best understood as a denial of a plausible-seeming identity condition for worlds, so I begin with some general remarks about such conditions. The weakest identity condition for worlds is what Fine calls *World Differentiation*:

> WD: if precisely the same objects stand in precisely the same relations at w_1 and w_2, then $w_1 = w_2$.

Nobody, I think, would deny WD: if worlds just are maximal states of affairs, and atomic states of affairs are adequately represented by n-tuples of objects and n-place atomic relations, then the same collection of atomic states will generate, by standard logical principles, the same collection of states (treating quantification as infinitary conjunction and disjunction). There would then be a single state which would obtain in virtue of that collection of states obtaining all together, and that single state is a world. There is no degree of freedom in this process which could somehow result in two different maximal states corresponding to the obtaining of the family of states generated by logic from a fixed collection of atomic states.

And there is nothing more to our conception of a world than that it is a maximal state of the sort indicated. So WD is irresistible (though its analogue for moments of time is not).

There is a stronger identity condition than WD which at least initially, seems equally irresistible. But it is this principle which Salmon controverts. Say that w_1 and w_2 are *qualitatively indistinguishable* iff there is an isomorphism from the domain of one onto the domain of the other (I assume that the behavior of the function on non-existents makes no difference). Suppose also that we divide the domain of a world into two classes, C_1, *constituted* objects, and C_2, *constituting* objects. For example, if a table A is constituted of a sum of wood W_A, A is a constituted object and W_A is a constituting one. More generally, a constituted object is a thing made of something, a constituting object makes something up (one object could be both, but for simplicity I ignore this). Then given a description **d** of a collection of states of affairs and the complete facts for **d** about which members of C_1 are constituted by which members of C_2, we can derive from **d** a description d′ at the constituting level by replacing names of elements of C_1 by descriptions of the form 'the thing made of α', where 'α' stands for the relevant element of C_2. With this apparatus, we can state an identity condition for worlds which I call Constitutional Sufficiency:[3]

> CS: if w_1 and w_2 are maximal states of affairs (worlds) which satisfy the same description at the constituting level, and w_1 and w_2 are qualitatively indistinguishable, then w_1 and w_2 are numerically identical.

Note that CS does not say that there are no non-trivially isomorphic but distinct worlds: it does not identify worlds which are *simply* qualitatively indistinguishable. What it implies is that if an isomorphism between worlds is the identity map at the *constituting* level, it must also be identity at the constituted level. This is a very plausible sufficient condition for identity of worlds. To use an example I have employed before, refuting the condition would require us to make sense of something along the lines of the following: that there is a world like the actual world in absolutely every respect except that the steel tower in it in Paris opposite the Palais du Chaillot is a different tower from the one that is actually there. The tower, of course, is to be built of exactly the same steel, by the same designer, at the same time, etc., as the actual tower. But it is not to be the Eiffel Tower. Can the claim that there is

a world with such a tower be defended? I think not, but defending it is exactly Salmon's enterprise.

2. SALMON'S COUNTEREXAMPLE

Salmon holds that CS is vitiated by its failure to take into account the relation of relative possibility between worlds. He maintains the principle (a): a table A, actually made without leftovers from a sum of wood W_A, could instead have been made without leftovers from a sum of wood S such that S almost wholly overlaps W_A and conversely. So there is a world w possible relative to the actual world w* in which A is made of such a sum S. But principle (a) is evidently an instance of the following general thesis with the propositional form $\Box(P \supset \Diamond Q)$: necessarily, a table made without leftovers from a sum of wood X could instead have been made without leftovers from an almost wholly overlapping sum Y. But the truth of the embedded conditional here does not turn on some feature of the actual world: the conditional is a truth, indeed a necessary truth, at every world. So applying the conditional to w, we get a world w′ possible relative to w where A is made of a sum of wood S′ such that S′ almost wholly overlaps S and conversely. Iterating this gives us a sequence of worlds w* (the actual world), w, w′, w′′ etc., in which each world is possible relative to the worlds adjacent to it; and it gives us a sequence of sums of wood W_A, S, S′, S′′ etc. which constitute A in the correlated world. We can choose the sequence of worlds so that in its last member w_n, A is made of a sum of wood S_n which is completely disjoint from W_A.[4] Since Salmon also accepts the *necessity of overlap*, which is the principle (b) that A could not have been made from *entirely* different wood, his conclusion is that the relative possibility relation is nontransitive: w_n is not possible relative to w*, i.e. not a *possible* world, given the way things actually are.[5] The counterexample to CS arises because there is evidently a world u *possible* relative to w* where A does not exist, *some* table is made of S_n and things are otherwise as in w_n. u is indistinguishable at the constituting level from w_n, but is a different world nevertheless since it does not contain A, but in its place a different table B.

Salmon's position depends on drawing a sharp distinction between worlds, which he glosses as *maximal ways for things to be*, and *possible worlds*, which are that subclass of maximal ways for things to be comprising just the maximal ways things could have been.[6] His view is

that there is a way for things to be in which our table A is made from totally different wood S_n, but such *a way for things to be* is not *a way things could have been*. In general, there are hardly any constraints on the contents of ways for things to be: any formally consistent assignment of n-tuples of objects to n-place relations will count, even inconsistent ones (*op. cit.* p. 14). Thus, to illustrate with his own example, there is a way for things to be in which Nathan Salmon is a Visa credit card account with the Bank of America.

Salmon could agree with CS if it were understood as restricted to *possible* worlds, and he can point out that the intuition about tables which defenders of the necessity of overlap usually cite, that a table *could not* have originated from completely different wood, is upheld by his view, since worlds where A is made of S_n are not ways things could have been. But for all that it saves certain modal intuitions, I still regard Salmon's ontology of worlds as incoherent, and CS as correct in the unrestricted form in which I stated it. I will now try to explain why.

We can distinguish two ways of thinking of worlds, the *substantial* way and the *formal* way. Formal accounts of worlds are familiar in the literature: worlds as sets of sentences, or sets of propositions, or set-theoretic constructs of objects and properties (proxies for states of affairs), are all formalist theories. But there is some conception, or notion, of *which* they are theories, and this is the notion of a world thought of in the substantial way. It is this notion to which I intend to give voice in the doctrine that worlds are maximal states of affairs. Now there is no doubt that Salmon's idea of a world as a maximal way for things to be, subject to the minimal constraints that he imposes on it, is perfectly coherent when we think of worlds in the formal way. The two worlds u and w_n in his counterexample to CS are clearly distinct worlds, formally, since they are distinct sets of sentences, or propositions, or set-theoretic proxies for states of affairs, or whatever. But it seems to me that they are not distinct substantially. Furthermore, the formal approaches are trying to capture the intuitive idea of a world, and the substantial way of thinking is the intuitive way. So if I am right, Salmon would have made no case at the most important level for the range of worlds he wants to admit.

What is the formal/substantial distinction? The state of affairs of Socrates being a foot soldier has as its set-theoretic proxy the ordered pair <Socrates, being a foot soldier>. When we consider the state of affairs of his being a foot soldier, we are considering something that is just like

the state of affairs of his being a philosopher, except that it involves a different property and it does not obtain. But in the case of neither of these states of affairs are we considering an ordered pair. The ordered pair, one might say, represents the *content* of the state of affairs, while the state of affairs itself is what we can call the *concretization* of that content: in the concretization, the second member of the pair (the property) *inheres* in the first member of the pair (the object). Intuitively, in considering the state of affairs of Socrates being a foot soldier, we are thinking of the concretization of the content <Socrates, being a foot soldier>, that is, of the inhering in Socrates of the property of being a foot-soldier, and not of some ordered pair of entities.

This may make it seem that to think of the concretization of a state-content is to think of that content being the content of a state which actually obtains, in which case the formal way of thinking would be the only appropriate one for possible but non-actual states. But this is not what I intend by the notion of concretization. Within the philosophy of modality there is a distinction between *modalists*, for whom modal operators are the primitive means of expression of modal propositions, and *anti-modalists*, who hold that modal operators are to be understood as quantifiers over possible worlds (which, on the most common view, are actually existing abstract objects). But a modalist may well have possible worlds *within his ontology*, even if he does not use them to give a semantics for $\Box$ and $\Diamond$. Again this position divides: a *possibilist* will have the actual *and* non-actual worlds within his ontology, while an actualist will have only the actual world within it. Indeed, the strongest actualist argument for an ontology of worlds is that the actual world exists, though this is not quite trivial, since it presupposes the ontological legitimacy of a certain kind of completed totality. [In fact, the anti-modalist position subdivides in the same way, to give a total of four possible combinations of views.]

A modalist-actualist (henceforth 'modalist' – I will ignore the modalism-possibilism combination[7]) can express facts about non-actual worlds using modal operators, since each non-actual world exists at itself (and at no other world, by actualism). Thus for 'P fails at some possible world' the modalist can say '$\Diamond(\exists x)(Wx \ \& \sim Px)$', where 'Wx' means 'x is a world' (note: not '*possible* world' – possibility is guaranteed by the $\Diamond$).[8] A modalist need not deny that there are actually existing abstract objects which, for purely heuristic purposes, may be treated as the domain of quantifiers which render modal operators. But it is completely obscure

how such extensional translations could *explain* the meanings of modal statements.[9] Rather, which such abstract objects should be counted as worlds will depend on which possibility and necessity statements are true. Any ontology of *abstracta* not conforming to this constraint is simply irrelevant to the metaphysics of modality. Thus the anti-modalist's '$(\exists\ w) \sim Pw$' is true or false in virtue of the truth or falsity of '$\Diamond(\exists\ x)(Wx\ \&\ \sim Px)$'.[10] Hence Salmon's thesis that there are impossible but possibly possibly...possible worlds where A is made of S_n is true only if some statement of the form '$\Diamond\ldots\Diamond(\exists\ x)(Wx\ \&$ Made-of-at $(A, S_n, x))$' is true, for sufficiently many $\Diamond$'s. The notion of the concretization of a state-content is the notion of a *truth-maker* for a statement of this form, specifically that part of the truth-maker which answers to 'Made-of-at (A, S_n, x)'. That is, Salmon's apparatus is acceptable only if we can make genuine sense of the '$\Diamond\ldots\Diamond$'-statements whose truth is required for the actual existence of the ways for things to be, *if* quantification over these ways for things to be is to be an extensionally correct rendering of modal operators (again, I do not deny that the ordered pair <Salmon, being a Visa account at BA> exists – the question is whether its existence is relevant to any issue in the philosophy of modality). And what is it to have a conception of a truthmaker for such a statement as '$\Diamond^n$ (A is made of S_n)'? Here as elsewhere, such an ability is presumably structured: we combine our understanding of what it is for something to be possible or possibly possible or... and so on, with our conception of A's being made of S_n. And my claim is that we have no such conception.

Let us return to Salmon's controversial candidate for a way for things to be. To make the example more vivid, suppose that in the series of worlds $w^*, w, w', w'', \ldots, w_n$, the relevant artefact in the last world w_n is not even a table or made of wood: as its original matter is bit-by-bit altered from world to world in the sequence, plastic is substituted and the form is also changed slightly, so that in w_n we arrive at a plastic chair. Now it seems to me that there is a considerable difference between the following:

(1) Concerning the wooden table A, consider a state of affairs in which that table is a plastic chair. Could such a state of affairs obtain?

(2) Consider the state-of-affairs content <A, being a plastic chair>, where A is this wooden table. Is it possible for this content to be the content of a state which obtains?

Salmon's idea is that we can perform the mental feat requested in (1) and then bring essentialist principles to bear on the state of affairs thus grasped, to classify it as possible or impossible. But I think that we can only perform the mental feat requested in (2). We can think of the ordered pair <A, being a plastic chair>, where A is a wooden table, since this pair is just a certain set whose ur-elements are an unproblematic object and an unproblematic property. However, if we try to think in the substantial way of a state of affairs in which A, the actual table, is a plastic chair – a state in which *being a plastic chair* inheres in A – it seems to me that we collide with a barrier of intelligibility. My explanation of this difference between (1) and (2) is just that in the substantial way of thinking of states of affairs, we are thinking of concretizations; and the concretization of the content <A, being a plastic chair>, escapes our grasp. Correspondingly, the notion of a way for things to be of which this concretization is a part escapes our grasp.

Why is the concretization of <A, being a plastic chair> beyond our grasp? I suggest that the reason is as follows. We cannot think of the two states of affairs of A being a wooden table and of A being a plastic chair, or of Nathan Salmon being a human being and Nathan Salmon being a credit card account, unless we have the conception of a single thing such that one state of affairs consists in that thing having a certain range of intrinsic properties, and the other consists in that same thing's having a radically different range. But no concept of ours answers to the notion of 'thing' that would be required here. For such a notion would be the notion of an object as being constituted by an overlay of properties upon a characterless substratum or *bare particular*, and I think the idea of a bare particular is of dubious intelligibility outside the realm of mathematical abstraction. Supposedly, we could arrive at such an idea by 'removing in thought' all the intrinsic properties of an object. But there is no such process. Any given object is an object of some specific ultimate kind, and if we remove the kind property, as the examples which Salmon willingly embraces require, what we are left to contemplate is not a characterless substratum, but rather, nothing at all.[11]

I reject Salmon's ontology of ways for things to be, then, because I think it is committed to the conception of the relation between subject and predicate as a relation between a bare particular and a quality that inheres in it. There is nothing question-beggingly modal about this objection. I am not saying merely that it is not possible for there to be a state of affairs in which Salmon is a credit card account or A a plastic chair. I am saying

that the conception of such a state of affairs is unintelligible since it presupposes the conception of characterless substratum, which is itself unintelligible. Thus no way for things to be – much less any way things could have been – answers to the state of affairs Salmon purports to describe. And CS is not defeated by 'counterexamples' which depend on taking modal operators to be quantifiers over entities which are either literally inconceivable – if we really attempt to use concretizations – or no more relevant to the truth or falsity of modal sentences than whatever counterexamples could be generated by taking these operators to be 'quantifiers over towns restricted by the relation of being connected by rail'.[12]

3. THE BRANCHING CONCEPTION

We stated the principle CS in the broadest way, as a principle about worlds, not merely possible worlds. But if the conclusion of the previous section is correct, we have no reason to think anything is to be gained from maintaining this distinction. In what follows, therefore, 'world' and 'possible world' will be used interchangeably. I wish to end by considering one further identity principle for worlds, which makes a stronger claim than either WD or CS. The principle appears to be false, so if appearances do not deceive, we have identified a boundary beyond which fineness of discrimination of worlds is carried too far.

The principle in question is:

> Q: If there is any isomorphism between **w** and **u** at all, then **w** and **u** are identical.

Q says that qualitatively indiscernible worlds are identical, and as was noted in §1, is stronger than CS. That Q is implausible is brought out by a well-known example of Robert Adams'.[13] Let w be a world where there are two qualitatively indiscernible and eternally existing globes. Suppose that these globes are only contingently future-immortal, and that at any time t, either could cease to exist. Then there is a world u just like w except that at a fixed time t one of the globes ceases to exist, and a world v just like w except that at the same time t the other of the globes ceases to exist. Call the globes α and β. Then there is an isomorphism between u and v which maps α to β and which, under reasonable assumptions, is a linear map of the entire space of u onto the entire space of v. u and v constitute a counterexample to Q.[14]

Is it just a curiosity that our intuitions about possibility and necessity support CS while failing to support Q?[15] Or is there some underlying conception of the nature of possible worlds which we employ in thought, which implies CS but not Q? I think there is such a conception, the *branching* conception. My version of this conception is as follows. We can think of a world as a continuous sequence of 'total' states of affairs, one for each moment of time, a total state for t consisting in the conjunction of all the states which obtain at t. On the branching conception, any two worlds which have objects in common consist in two sequences of total states with an initial segment in common, as in a 'Y' configuration (but possibly with an infinite descender etc.). Some point in time is the first point after which different states obtain at the worlds, and that is the branch point between them. *Note that there is no requirement that any objects in common must begin to exist before the branch point.* There can be transworld identities between objects which come into existence after the branch point, if these identities are fixed by facts about the common initial segment. Thus if a sperm and an egg exist in the common initial segment of two worlds w and w′, and these worlds branch before the sperm fertilizes the egg, we can still identify the zygotes that result in w and w′ when the sperm does fertilize the egg in each world, since by reference to the common initial segment, we note that it is the *same* sperm and egg in w and w′. Similarly, we can transworld identify tables which do not exist until after the branch point, in part for the reason that they come from the same wood, wood which existed in the common initial segment of the worlds. Obviously, the procedure can be extended to objects well after the branch point.[16]

One might wonder whether the view that there are worlds which differ only by a permutation of individuals in the manner of Adams' example requires an ontology of bare particulars, even though such a pair of worlds does not by itself refute CS (because the permutation goes all the way down). But the branching picture brings out exactly why Adams' example requires nothing so mysterious to explain its (undeniable) intelligibility. Moreover, in the absence of bare particulars, the branching conception implies CS, since it is only on the assumption that the identity of a thing does not supervene on the identity of the objects and processes involved in its coming into existence that a *branching* counterexample to CS can arise. It seems, then, that we have found a unifying conception which explains the status of both CS and Q.

However, this would be a poor explanation of the truth of CS and the

falsity of Q if the branching conception is independently implausible or not conceptually prior to these principles. I conclude this paper by defending its plausibility and priority.

The threat of implausibility arises from the fact that branching conceptions of worlds have consequences for the essential properties of individuals. Thus a version of the conception on which all common objects must come into existence at or before the branch point will make all the circumstances of a thing's origin essential to it, as well as every detail of the history of the world up to the moment of the thing's origin. This violates clear intuitions of contingency: surely the prototype of the Apple I (that particular physical machine) could have been built in Jobs' garage instead of Wozniak's? But my version of the branching conception also has an essentialist consequence: it is essential to an object that not everything about the history of the universe before it comes into existence be different. More precisely,

$$\text{(3)} \qquad \Box\,(\forall x)(x \text{ actually exists} \supset (\exists t)(\forall t' \leq t)(\forall \alpha)$$
$$(\alpha \text{ is the total state at } t'$$
$$\alpha \text{ is actually the total state at } t')).$$

Does this violate our modal intuitions? Perhaps I could have existed even if it had been that for some past moment, at each moment previous to it at least one thing's shape is slightly different from its actual shape. But I think it is both difficult to decide whether we do think this is possible (contrast the Apple I) and unclear that it is inconsistent with the present version of the branching conception: perhaps everything material could have been different, for some world branches from the actual world before any matter comes into being. Further search for counterexamples, involving e.g. spacetime structure, takes us yet further from cases where we know what the answer should be.

However, there is one much more threatening case. If it is contingent that the two globes in Adams' example are eternal – if one might have ceased to exist while the other continued to exist – surely it is also contingent that either globe exists at all. Yet a world where only one of the globes (ever) exists violates the branching conception. It was to accommodate this case that in *The Metaphysics of Modality* I allowed an exception to (3). We can imagine that a possible world w consists of 'isolated halves' (it does not matter for my discussion how isolation is to be construed, so long as it is sufficiently strong). In that case, perhaps there are worlds u and v consisting in just the respective halves of w.

Adams' world w, a world u where only one globe exists, and an indiscernible v where only the other globe exists, satisfy this description if the globes are isolated in w. Stephen Yablo has raised the question whether admission of u and v is consistent with the requirement that is the main thesis of *The Metaphysics of Modality*, that transworld identity and diversity should be intrinsically grounded.[17] Since u and v have no common initial segment, what *makes it the case* that they contain different globes?

With regard to ordinary branching in the 'Y' configuration, I deny that genuine transworld identity questions arise about objects in the common initial segment, since it is numerically the same segment of a course of events that constitutes that part of the history of the branching worlds. I believe that the same can be said of worlds which can be conceived of by removing part of a given world and making no other changes in what happens. Instead of the 'Y' configuration, we have the following picture:

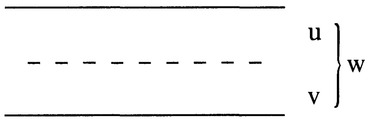


Figure 1

Here w is a single world, u and v its isolated halves.[18] [It may seem from the picture that three worlds exist at w, but this is incorrect. For example, at w it is false that there exists a world where only the events constituting the 'top' layer of w occur, though prefixing this claim with '$\Diamond$' yields something true at w.]

My claim is that just as there are no substantial transworld identity questions about objects in the common initial segment of a pair of branching worlds, so no substantial questions arise about identity amongst the elements of the domains of u and v. What makes the globes of u and v distinct is their distinctness in w together with the relationship between w, u and v illustrated in Figure 1, a relationship I call 'branching' in an extended use of the term. Yablo suggests that to say that the globes of u and v are distinct because there is a world like w in which the globes stand apart 'is arguably to get the intuitive order of explanation reversed: it is because the globes are distinct that they can stand apart' (*op. cit.* p.334). I agree that it is because they are distinct that they stand apart in *w*. But then the numerical differences between the indiscernible worlds u

and v are grounded in the distinctness of the two halves of w, which is itself based on the *intraworld* numerical distinctness of the globes.

However, it can reasonably be objected (Yablo did object) that this response gives w a special status which it does not merit. For w is not the only world from which u and v branch in which the globes are distinct: there is no limit to the number of other worlds w ′, w ′ ′, w ′ ′ ′ somewhat like w from which u and v branch. So it seems that I have to say that what grounds the distinctness of the u-globe and the v-globe is that they are distinct in w, w ′, w ′ ′, w ′ ′ ′, which again may seem to put the cart before the horse: aren't the branching relations in which u and v stand to w, w ′, w ′ ′, w ′ ′ ′ to be explained by the fact that the u-globe and the v-globe are different globes, rather than the distinctness of the globes being explained by the branching relations in which u and v stand to w, w ′, w ′ ′, w ′ ′ ′? On the former view, we use one identity fact to explain many branching facts, while on the latter, an identity fact seems to be grounded by any or all of a number of branching facts, no one of which of could be *the* ground of the identity; and if we say that all together they constitute the ground, we certainly invite the charge of putting things the wrong way round.

I am unpersuaded by this objection. First, we have to consider what the worlds w ′, w ′ ′, w ′ ′ ′ are like. If u and v are both to be parts of w ′ then they must be isolated in w ′, else w ′ would have to differ from w with respect to its portions which constitute u or v (so in fact u or v would *not* be a part of w ′). Hence w‘ must differ from w simply in having *more* parts; for example, perhaps u and y are isolated *thirds* of w ′. So w stands to w ′ as u and v stand to w: we can think of w as obtained from w ′ (u and v as obtained from w) or of w ′ as obtained from w by adding parts to it (w as obtained by merging u and v). Hence all the worlds w, w ′, w ′ ′, w ′ ′ ′ stand in the same sort of problematic relation to each other as u and v do to w: transworld identity facts amongst them are not settled by intraworld identity facts concerning a common initial segment.

But I do not agree that the transworld identities are ungrounded by the branching relations. To hold this may be to take an unacceptably realistic view of possible worlds. In Adams' case, we are given w by stipulation. *Granted* that w has isolated parts, we may then introduce u by the stipulation that it is constituted by one of these parts and v by the stipulation that it is constituted by the other. In this sense, the possibility of distinct worlds u and v turns on the distinctness of the globes, but that numerical difference derives from the globes being given to us in a single world. Suppose we try to proceed in the other direction: we simply

stipulate that there is a world which contains a globe α and a qualitatively indiscernible world which contains a different globe β, and announce this as a counterexample to Q. It seems to me that nothing is accomplished by this. The defender of Q will ask what makes the stipulated worlds two rather than the same world twice over – after all, not just any stipulation with implications for what is metaphysically possible is acceptable – and such a defender can raise the same question for each answer she is given. If the worlds are distinct because the globes are distinct, what makes the globes distinct? If it is because they are made of different matter, the question of what grounds this difference then arises, and so on: a sceptic such as Field about the distinction between u and v would not be moved by such a strategy.[19]

On the other hand, if the world w is given from the outset, then the counterexample to Q is conclusive. Of course, it is equally conclusive given any of the worlds w′, w′′, w′′′ from the outset, but this does not mean that the branching relations in which u and v stand to any one of w, w′, w′′, w′′′ are insufficient to ground the distinctness of the globes in u and v. It merely means that any triple of worlds consisting in u, v and one of w, w′, w′′, w′′′ has the feature that grounds the distinctness of the globes of u and v, and it is hard to see any other candidate for this feature than that a single world is given in which there are incontestably distinct globes, and u and v consist, by definition, in different portions of this world, each portion encompassing its own globe. So the intraworld distinctness of the globes really is prior to the distinctness of the globes of u and v.

Earlier I said that an identity or non-identity between x in w_1 and y in w_2 is not a genuine *transworld* identity or non-identity if it is settled by an initial segment of events common to w_1 and w_2. The difference between the globes of u and v is also settled by a numerical distinction which arises within a single course of events, so I think we have equally good justification for regarding it as not genuinely a transworld non-identity. In that case, we can admit the kind of exception to (3) generated by a world with isolated parts without reneging on a commitment to making only those genuinely transworld numerical identifications and discriminations which can be grounded in the intrinsic features of the things identified or discriminated.[20]

NOTES

[1] See Salmon, 'Modal Paradox: Parts and Counterparts, Points and Counterpoints', in *Midwest Studies in Philosophy XI: Studies in Essentialism*, French et al (eds.), Minnesota University Press 1986, pp. 75-120, and 'The Logic of What Might Have Been', *The Philosophical Review* 98 (1989) 3-34.

[2] I give a formal theory of states of affairs in my *Languages of Possibility*, Oxford: Basil Blackwell 1988, Ch.5. A simplified version is to be found in my 'Truth, Correspondence and Redundancy', in *Fact, Science and Morality: Essays on Language, Truth and Logic*, Macdonald, G. and Wright, C. (eds.) Oxford: Basil Blackwell 1986.

[3] I take this nomenclature from the title of the paper 'Against Constitutional Sufficiency Principles' by T. J. McKay, in French *et al, op. cit.* pp. 295-304. I hope to discuss McKay's challenge to CS elsewhere.

[4] The example is originally from 'Identity Through Possible Worlds: Some Questions' by Roderick Chisholm, *Noûs* 1 (1968) 1-8.

[5] This proposal is sensitive to the way in which the problematic premisses are formulated. In Salmon's exposition, as in Chisholm's, the premisses are given the form $\Box(P \supset \Diamond Q)$, which, together with $\Diamond P$, entails $\Diamond Q$ only if accessibility is transitive (otherwise all we may infer is $\Diamond\Diamond Q$). But the fundamental intuition underlying Chisholm's Paradox is that a sufficiently small amount of change cannot make the difference between a world representing that x exists and representing that x does not exist. This intuition is more naturally formulated along the lines of: if such-and-such a constitution is possible for x then so is such-and-such, in other words, as a conditional of the form $\Diamond P \supset \Diamond Q$, which with $\Diamond P$ *does* entail $\Diamond Q$ in classical sentential logic. So on Salmon's approach, one version of the paradox fails because a principle of modal reasoning is incorrect ($\Diamond\Diamond A$ does not yield $\Diamond A$), while if presented with the other version, Salmon must say that one of the premisses $\Diamond P \supset \Diamond Q$ is false. The alternative approach to Chisholm's Paradox for which I have argued (see *The Metaphysics of Modality*, 1985, Oxford, Ch.7) treats both versions in precisely the same way. And if, as I have also argued, Chisholm's Paradox is a Sorites paradox, the version with conditional premisses of the form $\Diamond P \supset \Diamond Q$ must be the fundamental one, and a resolution which tinkers with modal logic cannot be fundamental, since standard Sorites paradoxes involve no modal logic nor any underlying piece of machinery at all comparable to the relative possibility relation.

[6] See 'The Logic of What Might Have Been', especially Section III.

[7] I argue in Ch. 2 of *Languages of Possibility* that the best motivations for possibilism, if accepted by a modalist, would inevitably drive her towards anti-modalism.

[8] For further discussion of modalism and actualism, see Part I of my *Languages of Possibility*. The identity conditions for possible worlds formulated at the beginning of this paper were expressed in anti-modalist fashion (on a modalist-actualist reading, they are all trivially true, since there is only one world for their quantifiers to range over). Modalist versions are somewhat unwieldy, but formulable, and also, on my view, more fundamental. For example, the modalist version of WD is: $\Box_1 \forall w \Box_2 \forall w' \Box x \Box \forall P[(@_1 Px \equiv @_2 Px) \supset w = w']$ (see *op. cit.* Ch. 4 for notation).

[9] See Ch. 4 of *Languages of Possibility* for a defence of this remark.

[10] This kind of reduction of the extensional to the intensional was pioneered by Kit Fine. See e.g. his 'Plantinga on the Reduction of Possibilist Discourse' in *Alvin Plantinga*, edited by

James Tomberlin and Peter van Inwagen, Reidel 1985.

[11] Here I generally follow Wiggins, though he would be dubious about the application of his apparatus to artefacts such as chairs. See his *Sameness and Substance*, Basil Blackwell 1985, Ch. 5

[12] See *On The Plurality of Worlds* by David Lewis, Basil Blackwell 1986, p.20.

[13] See his 'Primitive Thisness and Primitive Identity', *The Journal of Philosophy* 76 (1979) 5-26.

[14] In his review of *The Metaphysics of Modality* in *The Philosophical Review* 97 (1988) Phillip Bricker ascribes to me the view that 'if a exists at a world w and w is qualitatively identical with a world v then a exists at v' (p. 129). But my endorsement of Adams' example there shows that I do not accept this principle. The principle relevant to the example which Bricker cites in support of his attribution is CS, not Q.

[15] Adams' example is consistent with CS since the globes of u and v are constituted of different iron.

[16] This is my response to David Kaplan's remark that 'for individuals not extant during an overlap...tracing back to the common part and comparing...[is] unavailing'. See his 'Quantifying In', in *Words and Objections*, Davidson, D. and Hintikka, J. (eds.), Reidel 1969, at p.224.

[17] See his review of *The Metaphysics of Modality*, in *The Journal of Philosophy* 85 (1988) 329-37.

[18] No great changes are needed in (3) to make it consistent with Figure 1: given a criterion of isolation, the total states quantified over can be restricted to states which do not draw upon isolated parts.

[19] See 'Can We Dispense with Space-Time?' by Hartry Field, in *Proceedings of the 1984 Biennial Meeting of the Philosophy of Science Association* edited by P. Asquith and P. Kitcher, fn. 14.

[20] In writing this paper, I have benefited from discussion and correspondence with Jonathan Lowe, Nathan Salmon, Jim Stone and Stephen Yablo.

MARCO SANTAMBROGIO

WAS FREGE RIGHT ABOUT VARIABLE OBJECTS?

Beginning with Frege, modern logic has completely changed our views on the perennial problem of universals. In fact, as traditionally understood, universals have almost completely disappeared from logic. Their peculiar nature was mainly due to the fact that they were at the same time both subjects and predicates: it was thought that "homo" refers to the same entity, whatever that might be, both in "Socrates est homo" and in "Homo est animal". It is a well known thesis of Frege's that there is no entity which is capable of playing that twofold role: subjects and predicates are sharply distinguished, semantically as well as syntactically – or, to put it in Fregean terminology, no object could possibly be a concept and vice versa.

That thesis is of fundamental importance, of course, but it is not the only one that Frege puts forward concerning universals. In almost all of his papers devoted to the notion of a function, from "Function and Concept" and "On concept and object" published in 1891-2, up to "What is a function?" (1904) and elsewhere, Frege is eager to discredit the notion of an arbitrary, variable or indefinite object (he seems to take these three adjectives as basically equivalent). As a result of such criticism we have the modern notions of function and variable and the logic of generality based on the quantifier-variable notation. As a byproduct, we also have the view that individual terms can only stand for individual objects and not for objects which possess any measure of generality.

Attributing to the natural language sentence "Homo est animal" the logical form: $(x)(Hx \rightarrow Ax)$, evidently depends both on Frege's sharp distinction between objects and concepts and also on his thesis that there are no indefinite (arbitrary, variable,...) objects. These two premises are independent of one another. This is shown by the fact that it is possible,

after all, to give a respectable semantics for indefinite or arbitrary objects, referred to from within (a suitably extended) first-order language, which acknowledges the distinction between subjects and predicates. One such semantics has been given by K. Fine in his recent *Reasoning With Arbitrary Objects*; another one has been given by the present author[1]; neither of them is likely to be the final one. (Of course the theory of denotation and of variables put forward in Russell's *Principles of Mathematics* should also be mentioned in this connection, but it is far from acknowledging any ontological distinction between subjects and predicates)[2]. That such semantics are possible is not, in itself, a complete answer to Frege's arguments against the existence of arbitrary objects. These arguments have to be carefully considered in their own right. This I shall begin to do in this paper; the whole subject of arbitrary objects is so ramified and touches upon so many important topics (as the notion of arbitrariness is in any case an important one in mathematics), that many points will have to be left out.

The first of Frege's attacks against arbitrary objects can be found already in the *Grundlagen*. It does not amount to an argument, but it is well worth quoting the whole passage, for future reference:

> It is true that at first sight the proposition
>
> "All whales are mammals"
>
> seems to be not about concepts but about animals; but if we ask which animal then are we speaking of, we are unable to point to any one in particular... As a general principle, it is impossible to speak of an object without in some way designating or naming it; but the word "whale" is not the name of any individual creature. If it be replied that what we are speaking of is not, indeed, an individual definite object, but nevertheless an indefinite object, I suspect that "indefinite object" is only another term for concept, and a poor one at that, being self-contradictory.[3]

For Frege, "all horses are four-legged animals" is in fact what is expressed by "the horse is a four-legged animal"[4]; similarly, "all whales are mammals" is really the same as "the whale is a mammal". This passage suggests, albeit inconclusively, some similarity between such "indefinite objects" and Meinongian "incomplete objects". Both kinds of putative entities, in any case, are to be sharply distinguished from the corresponding attributes: *the horse*, and *the whale*, as indefinite objects, are certainly distinct from the attributes of *horsehood* and *whalehood*. In fact, Frege had no reason for rejecting the latter. The nominalization that yields "horsehood" and "whalehood" cancels the predicative character of the predicates ("being a horse" or "x is a horse", and "x is a whale") to which

it was applied to give perfectly definite, if abstract, objects.[5]

The distinction to which I am pointing was made as a matter of course in traditional logic – the standard example being *album vs albedo*; it was generally referred to as that between the concrete and the abstract.[6] It is only the latter kind of entities (that I called attributes above) which has recently attracted some attention, so that we now have several theories of properties. But it seems to me that *the horse* and *album* are ontologically far more interesting than *horsehood* and *albedo*. For one thing, the properties of an abstract attribute such as *horsehood* or *albedo* are almost entirely disjoint from those that hold of the individuals satisfying the corresponding predicate: *horsehood*, unlike *the horse*, is not a horse, nor an animal; *albedo*, unlike *album*, is not white, much less a colour. It is for this reason that, when traditional logicians thought that indefinite entities were needed in order to prove general propositions concerning all individuals in a given class, it was the former kind of entity rather than attributes that they had in mind. E.g., in proving some theorem concerning all triangles (that the sum of their angles is 180 degrees, say), one is naturally led to consider *the* arbitrary *triangle*, which is a triangle, or triangular: if this satisfies the theorem, people used to think, then all individual triangles must also satisfy it.

Triangularity, on the other hand, is simply irrelevant in this connection, and it is not quite clear where else it could be of any use. Surely its *being in* a particular individual does not explain why that individual satisfies the corresponding predicate, i.e. why it is a triangle; on the contrary, the flow of explanation is, it seems likely, in the opposite direction. *Because* an individual satisfies a given predicate, we can say that the corresponding attribute *inheres* in it. It is not even quite clear that such entities can be taken as a modern substitute for traditional universals – as many philosophers have thought. It is not at all obvious, in particular, that being exemplified or instantiated by many individuals is reason enough to be taken as a universal. It seems to me that triangularity, which is instantiated by all triangles, is itself still a perfectly definite, individual entity. In order to approach somehow the traditional notion of a universal, instantiation by many is not enough: at the very least, satisfaction of the corresponding property is also required. *The* arbitrary or indefinite *triangle* (provided, of course, that there is such an entity) both is instantiated by all individual triangles and satisfies the predicate of being triangular. This is as near as we can come to the traditional conception of the universals as what can be at the same time both a subject and a predicate, given Frege's veto.

Frege's paper "What is a function?" is devoted to demolishing Czuber's notion of *variable* object. As the examples chosen by Frege will show, variable objects are simply indefinite objects renamed. The first argument given by Frege draws on the difference between variation and substitution.

As far as variable quantities are concerned, we must realize, says Frege, that in reality there is nothing that varies. One can indeed naively think that the expression "the number that gives the length of this rod in millimeters" names a number – a variable number, since the rod does not always keep its length constant. However, there is really no variation here: as time passes, the number that gives the length of the rod at a given moment is *replaced* by another number, that gives its length at a different moment.

Similarly, "the king of this kingdom" does not refer, absolutely speaking, that is, if the time is not specified (that "this kingdom" is indexical is here irrelevant; let it refer, in the present context, to the United Kingdom). For different instants, it refers to different persons, replacing one another over the years. It should be apparent now, incidentally, that variable and indefinite objects coincide: just as *the king of this kingdom* takes as inputs instants of time and yields persons as values, in the same way *the horse* can take, e.g., spatio-temporal locations and yield individual horses.

Now, replacement or substitution is not variation, says Frege. In order to have variation, there must be an object that varies. This in turn means that something in that object must remain unchanged, while something else changes. It is the unchanging part which allows us to say that, before and after the variation, we still have *the same object*, which underwent a variation. When, in a given process, we are unable to find a sufficiently stable element identifying a single object, we are forced to say that two objects are involved in that process, one at the beginning and the other one at the end, and that in the process the former has been *replaced* by the latter. The case of "the king of this kingdom" is precisely of this kind, as we are in general unable to find a common element, identifying a unique person, in the various kings – save in the case where there is in fact a single individual fulfilling that role. We therefore have to speak of replacement or substitution, and the expression "the king of this kingdom" cannot refer to a variable object.

How convincing is this argument? I do not intend to discuss the conception of change underlying it, but let us assume that the expression "the king of this kingdom" cannot refer to a (variable) object unless there

is something that remains unchanged throughout the process of variation. But of course there is something that does remain constant through the process: each one of the kings is just this, a king of this kingdom. Clearly this is not enough to single out an individual or a unique material object, but it does identify a role, an office or a type. Now, why should we assume that change always takes place within an individual?

Suppose that a given printer is equipped with several fonts of characters: Elite, Pica, Times,... Even if there is no one character which is common to any two such sets, isn't it natural to say that it is the Latin alphabet which varies here, that the different fonts are *variations* of it, so that we do have variation and not just replacement of one set by another?

But perhaps one might deny that the case of the Latin alphabet is entirely parallel to that of the king of this kingdom. In the latter, even if there is an entity which remains constant through the replacement of one king by another, namely the royal office, this entity is totally irrelevant to our problem: by the description "the king of this kingdom" we normally do not mean the royal office, but rather the person fulfilling that role. It seems to me that the contrast between these two things is entirely parallel to that pointed out above, between the "concrete" term *album* and the "abstract" term *albedo*. Notice that a similar distinction is also available with types. If we say that the type exemplified by this token: |||, is composed by concatenation from || and |, then we take it as having the same properties (or at least some of the properties) possessed by any one of its tokens; whereas, if we take it as, e.g., the class of its tokens, we must attribute categorially distinct properties to it. But the same distinction can also be found in the case of the Latin alphabet.

Of course, none of this by itself shows that there is a coherent notion of a type *as an object* in the first sense (which is strikingly similar to that of a variable object). But neither does Frege's argument, by itself, exclude that. Now consider the following: if there is nothing which is the king of *this kingdom* unless an instant of time is fixed, then there ought to exist no statement with a definite truth value, in which "the king of this kingdom" occurs as a term. This is not so, however. In fact, there are many such statements: e.g., if this kingdom is the United Kingdom, it is definitely true that the king of this kingdom can do no harm. The definite truth or falsity of such statements seem to force upon us the idea that there is something, whatever it might be, which they are about. Frege's view seems to be at odds here even with mathematical practice, as the following quotation from D. Scott shows:

Take the equation: x2 = x + 1. Whether this is true or false depends on x, and such equations (generally) define a whole class of solutions. We can of course in this case investigate by well known methods exactly which x make the equation true; but by only the most superficial knowledge of the laws of algebra, we can easily assert a *conditional* like: $x^2 = x + 1 \longrightarrow x^6 = 8x + 5$. Indeed, all the values of x^n can be simplified under the assumption that $x^2 = x + 1$. Passing to the many examples we are familiar with in several variables, we see that conditional equations may often be verified even when a complete analysis of the solution set corresponding to the hypothesis is lacking. The assumption is *used* [my emphasis, M.S.] *as if it were true* even though by itself it has no determinate truth value owing to the occurrence of parameters.[7]

Whether or not we are entitled to take the step, from the fact that such statements are used as having a definite truth value to the positing of a single object which they are about, need not concern us for the time being, as we are now only interested in Frege's arguments. But notice that it is precisely because Frege does not think that such definite statements exist, that he claims – as we have seen him doing in the *Grundlagen* – that "As a general principle, it is impossible to speak of an object without in some way designating or naming it". In the statement "The king of this kingdom can do no harm" we do seem to be saying something of *the king of this kingdom* and, by the same token, of each one of the persons who ever fulfilled that office, without naming any one of them.

Let us now consider the second of the arguments against variable objects put forward in "What is a function?". As it is given in the course of presenting the modern, wholly Fregean notion of a variable, one must carefully distinguish that part of Frege's argument which speaks in favour of such a notion (obviously nobody would dream of going back on that) from the part which is directed against variable objects. It is quite possible that the effort of clarifying the modern notions of variable and function made it difficult for Frege to appreciate what can be salvaged in Czuber's (and others') variable objects; perhaps the notions of a variable and of a variable object express coherent, if different, intuitions. Here is the relevant passage:

Do we not use 'x', 'y', 'z' to designate variable numbers? This way of speaking is certainly employed, but these letters are not proper names of variable numbers in the way that '2' and '3' are proper names of constant
numbers; for the numbers '2' and '3' differ in a specified way, but what is the difference between the variables that are said to be designated by 'x' and 'y'? We cannot say. We cannot specify what properties x has and what differing properties y has. ... Since we cannot conceive of each variable in its individual being, we cannot attach any proper names to variables.[8]

Now, how much weight does this carry? It is difficult to say. On the one hand, Frege here seems to be obviously right. Is it not obvious in fact that *the horse, the whale, the triangle* are unique, in the sense that there can be only one thing that represents all the individuals (horses, whales, triangles) in a given class, or species? And so it seems that when we employ letters such as 'x','y','z', having the same range, it cannot be variable objects that they refer to; rather, they must be proper variables, in their modern sense. Incidentally, this is a delicate point also for Fine's theory of arbitrary objects (as well as for Russell's theory of the variables).[9]

On the other hand, it seems to me that there is something unconvincing about Frege's claim that for two things to be distinct we must always be able to point to some definite difference between them. Here, it all depends on the kind of things and on the sort of definite difference one has in mind. Take two *presentations* of some object – two views occurring at a very short interval from exactly the same viewpoint, of a perfectly immobile object. I do not have to explain what presentations are; Frege himself seems to be perfectly familiar with such things, for he often refers to them. Those two presentations have exactly the same content – i.e. what they show is the same – even if, being two by hypothesis, they are distinct. The only difference we are able to point out in them, however, is that one has taken place before the other. Having occurred at a given instant is in fact a well defined property of a presentation but it seems to be rather "external" and it is not clear whether Frege would consider it sufficient that we specify such properties as the "differing properties" which are needed in order to "conceive of each variable in its individual being". In any case, if variable objects were to belong to the same category of entities as presentations, they could very well be distinguished using such "external" properties (even if we would normally be interested in the properties of their contents). That variable objects belong to this category is a possibility, but it is not a thesis we intend to defend here. Notice, however, that such a thesis would not contradict the similarity, which we pointed out above, between variable objects and types: there is a clear sense in which a type can be said to present each one of its tokens. On the other hand, it cannot be objected to our defence of variable objects against Frege's argument, that the presentations of an object are not themselves objects, and therefore in particular cannot be variable objects. Given Frege's own notion of an object, presentations certainly *are* objects, in that they have names in language.

In the same article, "What is a function?", Frege also gives a slightly

different argument against Czuber's thesis that variables such as 'x','y', 'z', name variable objects. Frege concedes that expressions like "the number *n*", where no particular number is meant by "*n*", do occur very often in mathematics. Would it not be quite natural to suppose that there is something that is thus named, even if it is not a definite number? Here is Frege's answer:

we do sometimes say 'the number *n*'. How is this possible? Such an expression must be considered in a context. Let us take an example. 'If the number *n* is even, then $\cos(n\ \pi) = 1$'. Here only the whole has a sense, not the antecedent by itself nor the consequent by itself. The question whether the number *n* is even cannot be answered; no more can the question whether $\cos n\ \pi = 1$. For an answer to be given, '*n*' would have to be the proper name of a number, and in that case this would necessarily be a definite one. We write the letter '*n*' in order to achieve generality. This presupposes that, if we replace it by the name of a number, both antecedent and consequent receive a sense.[10]

Clearly, there is some similarity between this point and the one which we just discussed. But there is also something new. On the one hand, it seems that, taking these words literally, we have here a flat denial of the compositionality principle. Instead of a conditional having the sense it has in virtue of the senses of its antecedent and consequent, it is the other way around: the parts derive their senses from the whole, as "the antecedent by itself" and "the consequent by itself" have no sense. There is another problem, too; the context provided by Frege for the expression "the number *n*" is not the only possible one. Such expressions typically occur also in *assumptions*: "Let the number *n* be such and such". Now, not only does it seem to be hard to deny that assumptions by themselves do have meaning, as they are just the starting point of an entire piece of reasoning, but it is also implausible to take assumptions as always being the antecedents of a universally quantified conditional reaching as far as the conclusion of the reasoning. The problem here is the same as the one Russell had to face in connection with the axiom of infinity. Let A be such an axiom, B any proposition proved using A; Russell thought of taking the conditional $A \rightarrow B$ as the theorem actually proved. He thought that by discharging A, the purity of logic could be preserved. However, in view of the fact that, if A is in fact false, both $A \rightarrow B$ and $A \rightarrow \neg B$ are true, this view means that we are no longer entitled to treat $\neg B$ as false, even when we have already proved B. On the other hand, if assumptions are not treated as antecedents of conditionals, then it is utterly implausible to treat any variables occurring in them as implicity bound by a quantifier: if you assume that the number x is such that $x^2 = x + 1$, you are neither

assuming that every x is such, nor simply that there is such an x. This point has also been recently made by van Fraassen, who makes the following interesting remark:

> Well, [when assuming $x + y = z$] what exactly are you supposing? I'll offer an answer to such questions. You are supposing a proposition, but it is a proposition which depends on certain parameters [x], [y], [z], which you may or may not identify as specific entities – you may or may not "fix their values". This proposition is a sort of generic proposition, "general" in the sense, derided by Berkeley, of Locke's general triangle.[11]

However, it is possible that we have been unfair to Frege. It is possible that his remarks on the expression "the number *n*" being devoid of sense outside its context, was really a remark about reference. Perhaps all he is saying is that we can no more say what the value of "*n*" is than we can say who is beating whom in "he beats it" – taking this sentence out of such a context as, say, "If Pedro owns a donkey, he beats it". Here we seem to be in the same situation we found ourselves in with the assumption "Let $x^2 = x+1$". Frege seems to say that the variables, just like the pronouns in "He beats it", cannot have a reference (and therefore the whole sentence cannot have a definite truth value) if no context is provided to complete the sentence. But it is not at all clear that we cannot simply *assume* that the sentence "He beats it" *is* true; then the pronouns will simply have to refer to objects which are unknown to us. There is nothing strange in this, however. What it means is that we have no answer to such questions as "Who is beating whom?"; but are we any better off when we know that "Pedro is beating someone" is true? Here, too, we do not know whom Pedro is beating.

We shall leave aside for the moment some other difficulties which Frege raises in this paper, which seem to be directed mainly against Czuber's sloppy statement of the theory of variable objects. Let us now consider instead Frege's own rendering of such sentences as "The whale is a mammal", which seemed, at first sight, to be about the indefinite whale.

As we all know, according to Frege "The whale is a mammal" translates into "For every x, if x is a whale, then x is a mammal". It is not quite clear what rules govern the translation. Clearly, expressions like "the such and such" – when the context makes it clear that this is not a proper definite description – do not always correspond to universally quantified phrases. E.g., "The cow feeds the calf in the stable" does *not* translate into "Every cow feeds every calf in every stable". But what is the rule here? The problem is more complicated than it appears at first sight.

Consider the following two sentences:

1. The Greek hates the Turk,

and

2. The dog loves the master.

They are syntactically quite similar – unlike the pair "A boy loves every girl" and "Every boy loves a girl" – and yet we understand that the sequences of quantifiers occurring in the standard translation into first order logic, are quite different. What is not so clear is the kind of rule we are applying here. However, some philosophers might think that this problem is of concern only for linguists, and I will not pause in order to counter this false impression.

Now there is an obvious price to pay if we insist on eliminating indefinite objects. We are so used to it by now, that we might fail to realize its import; but it is at least uneconomical to treat the copula as ambiguous in three ways: besides the sense of identity, there are what Frege calls the *subter* and the *sub* senses.[12] "Moby Dick is a whale" and "The whale is a mammal" (which are quite similar both intuitively and on the indefinite object view) translate into two utterly distinct standard first order formulae, the former being of the subject/predicate form, the latter being the subordination of one concept to another. Some linguists, notably Jackendoff, have recently objected to this as quite intolerable.[13]
Although it would be silly to dismiss this problem as pertaining only to linguistics, I shall, again, pass on.

Second, there is a test suggested by Frege himself, which shows that something is amiss in translating

3. The whale is a mammal,

as

4. For every x, if x is a whale then x is a mammal.

"If in the sentence 'all mammals are land-dwellers' the phrase 'all mammals' expressed the logical subject of the predicate *are land-dwellers*, then in order to negate the whole sentence we should have to negate the predicate: 'are not land-dwellers'. Instead, we must put the 'not' in front of 'all'; from which it follows that 'all' logically belongs with the predicate".[14] What Frege seems to be saying here is that in order to negate a whole sentence one must put the negation in front of the predicate, the

real logical predicate. The behaviour of negation is, then, to be taken as evidence of the whereabouts of the logical predicate. Now, the negation of (4) is (5) "Not for every x, if x is a whale then x is a mammal". This means something quite different from (6) "The whale is not a mammal" which is the negation of (3). The latter implies that no whale is a mammal. This shows, according to Frege's negation test, that "is a mammal" is in fact the real predicate, and "the whale" is the subject: a structure quite different from that exibited in (4). Moreover, there seems to be no natural way of negating (3) in such a way as to produce the sense attached to the negation of (4); "Not all whales are mammals" could hardly be said to derive from (3) by negation.

At this stage, variable, indefinite and arbitrary objects seem to be closely related, if not utterly identical. They cut across Frege's distinction between objects and functions, in that they are objects and yet they are variable: one could even say that they take entities of some kind as input and yield proper, definite individuals as output. The *king of this kingdom* example shows how this can be effected.

Now, we all know how Frege proposes to deal with variables and with functions; but what does he have to say about arbitrariness? Before answering this, let us see why the notion of arbitrariness is crucial not only for proving universally quantified statements (which is the mathematician's favorite occupation) but precisely in order to understand the very notion of quantification in Frege.

Frege conceives of a universally quantified sentence as meaning the same as an infinite conjunction. If the language has a name for every object in the domain, then we can take the universally quantified sentence as equivalent to the infinite conjunction of all its substitution instances; whereas, if the language does not have all the names needed, we can take the logical product of all the truth values of the result of applying the property, expressed by the matrix of the quantified sentence, to each one of the individuals in the domain. In general, both these constructions involve operations on infinite totalities. As it is well know, many philosophers have objected to this on the ground that we do not seem to have the capacity of actually surveying infinite totalities and it therefore is rather mysterious how we could possibly understand such operations. I am not convinced that this objection is entirely well taken. For at no place, according to this conception, is it required that we be able to actually carry out an infinite operation or process. All is required is that we should be able to *imagine* the carrying out of an infinite task, and I do

not see that our *actual* limitations can be of any impediment to this process in *imagination*.

However, this is not to say that everything is beyond dispute in this conception. In particular, in taking a general statement as equivalent in meaning with an infinite conjunction, it seems that we are subscribing to a "distributive" conception of generality, so to speak. For instance, according to this conception, when asserting that man is mortal, we seem to be saying that this man is mortal, and that man is mortal and that other man is mortal, and so on and so forth (it does not really matter whether or not we are able to name each one of the men). But in saying that man is mortal, we do not seem to be thinking individually, or *distributively*, of any one man in particular, much less of every man individually. And part of the intuitive appeal that the notion of indefinite objects possessed can be explained by the fact that it seemed to offer a way of doing justice precisely to this intuition: if "man is mortal" is about the indefinite man, it does not (directly) concern any individual man in particular.

Now, the interesting thing is that Frege himself acknowledges that our saying that man is mortal does not involve a distributive conception of generality: he mocks the idea that in the thought expressed by "All men are mortal" (whether or not this is the same as "Man is mortal") we can find as a constituent the thought of the mortality of some African chief, of whom we have never heard.

> If I utter a sentence with the grammatical subject 'all men' I do *not* wish to make an assertion about some Central African chief wholly unknown to me.[15]

However, in a finite conjunction, and no doubt also in an infinite one, each conjunct *is* part of the thought expressed by the whole. So, how can the sentence "All men are mortal" have the form of an infinite conjunction? It is not the infinity of the conjunction which poses a problem here: it is rather the fact that, when asserting a universally quantified sentence, we seem to be asserting a conjunction, according to Frege, and yet we are not asserting any one of its conjuncts.

In his first book on Frege, Michael Dummett has given an explanation of this fact, which I am now going to restate as I understand it. First of all, in the case of a conjunction which is equivalent in meaning to a universally quantified statement, there is a clear uniformity which might be lacking in the ordinary case: each conjunct has the form of the matrix of the quantified sentence, with the free variables substituted for by the name of an element in the domain (at least, if the language has enough names to

name all such elements). It is therefore possible for us to detect a schema repeating itself in the infinite conjunction. Detecting it is the first step towards grasping the sense of the quantified sentence, without having to consider individually the sense of each conjunct, and therefore also the sense of each one of the names of the individuals substituted for the free variables (having to consider it would in fact make the thought of the mortality of the African chief part of the thought that man is mortal).

Once we have grasped the general form of all the conjuncts, we still have (a) to grasp the totality which constitutes the universe of quantification, since each conjunct arises by applying the matrix of the quantified sentence to each member in turn of that totality; and (b) to grasp what it means to say that the property expressed by the matrix of the quantified sentence applies to each member of the same totality. In both cases, the kind of grasp we are trying to achieve must be *general*: e.g., we cannot be required to grasp the totality of the domain of quantification by running through each one of its members, without immediately reverting to the "distributive" conception of generality. Similarly, we must arrive at a general idea of the set of all truth values resulting from applying the matrix of the quantified sentence to each one of the individuals in the domain, and therefore of its product, without having to consider individually each individual.[16]

Now, how are we able to grasp a totality *in general*? Dummett does not say, but we seem to have no choice left here but to take it as the grasping of some procedure, or a function, which determines for each object, whether or not it belongs to the totality in question. But of course, here again we must have a *general* grasp of such a procedure or a function: it would not do to take it as a set of ordered pairs which we must run through one by one, or we would be unable to overcome the distributive conception. Part of our problem is therefore seen to reduce to the well-known, but difficult problem of explaining what it is to grasp a function as a whole or "in general", without knowing which particular values it yields for which particular arguments. But let us leave this matter aside for a while and consider the second part of the problem, namely what it means to say that the property expressed by the matrix of the quantified sentence applies to any one of the objects in the given domain. It is here that the notion of an arbitrary object first comes in: we must know, in Dummett's own words, "what it is for the predicate to which the quantifier is attached to be true or false of any one arbitrary element of this domain".[17]

Obviously, what Dummett means here is not at all related to what Czuber had in mind; he means just the standard notion of an arbitrary object, which is employed throughout mathematics in proving general theorems. Let us briefly consider what it amounts to. It was Bishop Berkeley who stated most clearly, in his attack against Locke's general triangle (the most famous of all indefinite, arbitrary and variable objects), the conception underlying it. Berkeley thought that in order to prove a theorem about, say, all triangles, we do not have to consider the general triangle (of which we can only form, according to him, incoherent notions) as a separate entity, distinct from all the particular triangles. All we have to do, is to consider any one particular triangle and to prove that it has the property we are interested in. If the proof does not make use of the particular properties of that particular triangle (besides its being a triangle), then we can see that the proof can be repeated for any other particular triangle. That is, we can view the proof at hand as consisting of a schema with a slot to be filled in by the name of any one object in the domain. Seeing that the proof can be repeated for any other object just means that we see that any other name can be *substituted* for the given one, without the result ceasing to be a valid proof.

This conception of arbitrariness is precisely what is embodied in the Introduction Rule for the universal quantifier, which states precisely that, if we have a proof that a given property belongs to a particular object, about which nothing in particular was assumed, then we may conclude that every object in the domain has that property. There are a few remarks to be made. First, – as Dummett observed[18] – the meaning that such an introduction rule confers upon the universal quantifier is somehow more restricted than the intuitive conception we have of it. For the introduction rule requires that a *uniform* proof can be given for all objects in the domain, i.e., that the very same proof that was given for the particular object at hand can be repeated for any other, in order to assert the universally quantified sentence. However, it is conceivable that all objects in a given domain satisfy some property without there existing a unique form of proof applicable to all. Inductive proofs are in fact not uniform precisely in this sense.[19] It seems therefore that the intuitive meaning of the universal quantifier is not entirely exhausted by its introduction rule, or, in other terms, that the standard notion of an arbitrary object does not exactly match that of universal quantification.

This does not raise too serious a problem for the standard notion of arbitrariness. The case of another remark, due to Fine, is somewhat

different. Fine has observed that the standard account of arbitrary objects does not explain their use in proofs, e.g., by *reductio*. Suppose there are in fact no Fermat numbers (counterexamples to Fermat's theorem); still, it is legitimate to declare "Take an arbitrary Fermat number *n*" and then to go on and reason with *n*. But how can such an arbitrary object be an ordinary object chosen among the Fermat numbers, given that there are no such things?[20]

Finally, it should be noted that, according to this conception of arbitrariness, we do not possess any notion of an arbitrary object save with respect to some proof or other of a given proposition, since it is only in the course of giving a proof that some (perfectly definite and ordinary) individual is treated as arbitrary, i.e. as interchangeable with any other. The standard notion of arbitrariness is therefore a relative one: relative to (some proof of) some proposition or other. It is then open to us to appeal to arbitrary objects only when we have to convince ourselves of the truth of some *proposition*; but we cannot appeal to arbitrary objects if we are trying to grasp some non-propositional item.

We now have to consider whether this notion of arbitrariness is sufficient in order to explain how we have a "non-distributive" notion of generality. I think there is no problem in understanding what it means to say that the property expressed by the matrix of a universally quantified sentence applies to all individuals in the domain of quantification, given that it applies to an arbitrary one: here we are interested in saying that applying that property to any one individual in the domain yields the truth value True. It is therefore entirely legitimate to consider an arbitrary individual in the standard sense.

The problem then reduces to the following: can we have recourse to the standard notion of arbitrariness in order to grasp a function *in general*, i.e., without running through all its arguments and values individually or without knowing which individual values it yields for which individual arguments? (It will be remembered that grasping the domain of quantification *in general* was required in order to have a nondistributive conception of generality, and that reduced to the grasping of some procedure or function.)

How do we grasp any one function *in general*? This is closely related to the question, How can a finitist understand functions? In both cases, in fact, we are not allowed to appeal to a previous understanding of infinite totalities: in the case of the finitist, because he is supposed to be incapable of such understanding, and in our case because we are just trying to

explain what it is to grasp infinite totalities non-distributively. If we could appeal to some infinite totality previously understood, that understanding would clearly have to be non distributive; but if we already had it, we would not be bothering with the notion of a function. So, we are not allowed to assume infinite totalities at all.

The problem of explaining how a finitist can understand functions is a difficult one and I do not claim to have any new solution. Some, not all equally convincing, answers exist in the literature; I shall here consider William Tait's influential paper "Finitism".[21]

The first thing to be noted is that conceiving of a function as a rule or an algorithm is not of great help for our purposes. For, even if it is true that the rule of computation yields, as a matter of fact, a function, we must also be able to understand that the rule works. In particular, we must be able to see that it yields a (unique) value for each argument. Now, understanding this is to understand *inter alia* a statement of the form: (x)(Ey)G(x,y) – G(x,y) meaning that y is a computation of a value from x. But quantification is precisely what we were out to understand in the first place and so we have made little progress.

How can the finitist understand a function f from A to B (f:A —> B, for short)? Well, "he can understand it as recording the fact that he has given a specific procedure *for defining a B from an arbitrary A*".[22] This should not be taken as meaning that he understands the general proposition that for every x belonging to A, *fx* belongs to B, or we would be moving in a circle, or trapped in an infinite regress. The procedure or construction of a B from an arbitrary A must be of such a kind that "it makes no sense to ask for a proof, beyond the construction itself, of the corresponding proposition [(x:A)(fx:B) – i.e., for every x in A or, of type A, f(x) is in B or, of type B]."[23] Now, that there are in fact such constructions, rests on the fact that we can directly see in the way the arbitrary A and the arbitrary B are made or constructed that they allow certain constructions, or that they result from certain constructions: "to be a B is to be constructed in a certain way, possibly from an A. And [...] in virtue of the way an A is constructed it yields certain constructions, possibly of a B". E.g., writing C $\wedge$ D for the type of ordered pairs (c,d), with c in C and d in D , we can see that an arbitrary (c,d) consists of an arbitrary c and an arbitrary d, put together by pairing, and having seen this, by the same token and without any need for further proof, we should be able to see that pairing puts together in the same way any c and any d.

It is in this sense, according to Tait, that arbitrary objects allow even the

finitist to see how certain functions from one type to another (by no means all such functions) work, and therefore to understand them without reference to infinite totalities.

Our problem can now be restated thus: is the conception of arbitrariness which is being appealed to here, the standard one? That is, is the arbitrary a in type A any one *particular* object in A, which is interchangeable with any other object of the same type, in some given *schema*? (Notice that the problem of uniformity, which we met in connection with the universal quantifier, does not arise here: any function is uniform in that sense). Now, on the one hand, it seems that the answer is yes; at any rate, this is what Tait himself seems to believe: "Indeed, our terminology is otherwise misleading: as Frege pointed out, there is no such thing as an arbitrary A."[24] On the other hand, it is not entirely clear whether, having seen that a *particular* object in A can enter in a given construction, we are thereby also able to see that *all* other objects in A can do so, yielding similar results, as is required for Tait's argument to go through. In fact, it is quite possible that we see how some given, particular object enters in a given construction, without realizing that it is somehow arbitrary, so that the given construction has a wider significance and wider applicability. But what is involved in this further step of seeing a particular object as particular and, at the same time, as arbitrary? One wonders whether once again it all depends on some, possibly infinite, totality (e.g., of all the objects in the function's domain) being somehow grasped as a whole. If so, then no explanation has as yet been given of how the finitist can grasp a function or of how we can grasp a totality non distributively.

In any case, Tait's own elucidation of what an arbitrary object is in the case of the type N of natural numbers, seems to suggest a picture of arbitrariness which is somehow different from the standard one. *The* arbitrary number (for there seems to be only one arbitrary object for each type, according to Tait) is *the generic form of any finite sequence*. We discern finite sequences of objects of all kinds in our experience, and we see them as sequences – we see that they have some form in common, even if they are not all equally long; we see that they are all finite sequences. This form Tait calls *Number*. Just as with the forms of scholastic philosophy, there is an order among them: the form Number can be specified into various subforms, such as the even number, the square number, etc., and further down, the minimal such subforms being the particular numbers. 0 is the form of a null sequence, 1 of a (any one, of

course!) one-element sequence, etc. Those among us who do not like, for one reason or another, this scholastic talk of forms and subforms, can safely employ the type/token terminology here, which is at bottom equivalent. Forms are really types; e.g., 1 is the type whose tokens are all possible one-element sequences. Types have subtypes, of course.

Incidentally, the epistemological importance of such forms cannot be overrated: according to Tait, "[w]e do not understand Number via the concept of number, i.e., of being a number. Rather, it is the other way around. We understand the numbers as the specific determinations of Number". I shall not go into the reasons Tait gives for this claim, but notice that the situation here is the exact reverse of that obtaining in the case of attributes: it seems unlikely that the inhering of triangularity in a given object should explain why that object is triangular, whereas it seems plausible to hold that we understand forms first, and then we see that some object satisfies some predicate in that it has the corresponding form. But of course, this is much too general a topic to be discussed here.

Now, is this notion of an arbitrary object standard? First, forms are not objects, so that, given some domain of objects, the arbitrary representative of such domain cannot be an object in it. But, to this, one might reply that, according to Tait, particular numbers are in any case homogeneous with Number: they are all forms – the individual numbers being the minimal subforms of Number; so that they can be grouped together in one homogeneous domain. But even so, one cannot take Number, the generic form of a finite sequence (of every finite sequence) to be any particular number, 2, say, or 72 (*even if, of course, Number is a number, just as the triangle is triangular*).[25]

Second, when we discern the form Number in the finite sequences given to us in perceptual experience – e.g., when we see the following configuration: 111111111, as a sequence, even before counting the 1's – there is no particular proposition we are about to prove. But it will be remembered that, according to the standard view of arbitrariness, only relatively to the proof of some proposition or other does it make any sense to consider an object as arbitrary. This seems to show that Number is not an arbitrary object in the standard sense and it is rather closely related to the arbitrary, indefinite and variable objects that Frege disliked so much. This impression is reinforced by the remark that Tait's forms are types (in the type/token terminology) in that sense of the term, noted above, in which they share their properties with the tokens falling within them. In fact there is an interesting discussion in Tait's paper about what Hilbert

took the subject matter of "contentual" mathematics to be. Hilbert thought that mathematics is about the concrete signs and their syntactical relationships, but what is not entirely obvious is whether it is the sign token or the sign type which should count as concrete here. Tait offers several arguments for the latter, but the very fact that some uncertainly is left about what exactly Hilbert had in mind, shows that the types are understood here in the sense in which they share their properties with their own tokens.

So it seems that *the arbitrary A*, in Tait's sense, is not any one object chosen among the actual members of A. It then follows that his conception of arbitrariness cannot be the standard one. Must Tait then believe in pre-Fregean arbitrary or indefinite objects? Before deciding this, we have to answer a few crucial questions, such as: are the arbitrary (variable, indefinite) objects, as well as Tait's forms, *objects* at all? Are they *sui generis* (ideal) objects, to be added to the standard universe of ordinary objects? Do we talk *about* them? Ultimately, it is on the answers we give to these questions that our understanding of universal generalization rests.

The ancient doctrines about forms and substances are of some help here. Forms – they teach us – are not substances. Objects are rather like substances. Forms, therefore, and particularly Tait's Number, cannot be objects. We discern them in the objects and we see objects *as having them*; then, having discerned them in the objects, we can assert various things about the latter. For instance, having seen that some given arrangement of strokes has the form of a finite sequence, we can assert that one more stroke can be added to it in such a way that we still have a finite sequence. Thus, they give us *grounds for our assertions*. But they are not what we (normally) talk about. In a way, then, I think that Frege was right: there are no indefinite, variable or arbitrary objects. Number, in particular, is not an object; no wonder it cannot be found in the universe of numbers; nor does it make much sense to add it to the universe as a special or "ideal" object.

But is it not absurd to try to defend arbitrary or indefinite objects against Frege, and then to admit that Frege was right in saying that there are no such objects? I shall try to point to a possible way out of this embarassing dilemma. *The fact is that one does not have to defend the thesis that, e.g., the king of this kingdom is an object in order to take "the king of this kingdom" as a referring expression*. To my mind, the point of all theories of arbitrary, indefinite and variable objects (including

Meinong's) is not ontological inflation; their main point is rather that expressions such as "the triangle", "the whale", etc., should be taken as bona fide referring and not as a string of signs with no logical unity, to be paraphrased away. To refer is not necessarily to refer *to an object*, nor is what is referred to by a term in a given statement, what the statement is about. In fact, my claim is that expressions can refer to *forms*. As I said, forms are neither objects nor what we ordinarily talk about. What this claim amounts to is that we should distinguish two different notions of reference. (This is parallel to, and perhaps related with, the Meinongian idea of distinguishing two different kinds of existence. Actually, I am convinced that also within Meinong's theory one should draw the same distinction between two notions of reference). Clearly, there are a number of adjustments to be made with the other semantical notions, such as that of an object and, above all, of truth – as we shall now see. Also, one must show how the two notions of reference mesh.

How are we to understand then the statement "The king of this kingdom can do no harm", if "the king of this kingdom" is taken as referring to a form and not to an object or an individual? I think that one has to read that statement as saying that *being a king of this kingdom is reason enough for satisfying the predicate "can do no harm"*; the form *king of this kingdom* is what makes any individual, which has that form and insofar as *(qua)* it has it, satisfy the predicate. A form, unlike an individual, is something which can make a statement true or a predicate satisfied.

Take the statement "James can do no harm". What makes the predicate "can do no harm" satisfied by James? This is the same as asking: what makes it true that James can do no harm? One cannot answer: "James", for an individual is simply not the kind of thing that can be cited as a reason, explanation or ground for an assertion, although it can be what a statement is about. But one can in fact answer: "Its being king of this kingdom". Notice that, if one cites that as a reason, one is committed not only to the truth of "the king of this kingdom can do no harm" but also to the truth of any such statement as "Charles can do no harm", "George" can do no harm" etc., provided that Charles, George, etc. are kings of this kingdom. Whatever can be cited as a reason for satisfying a predicate must be *general*, in the sense exemplified here. It is simply not intelligible to assert that some individual is P *in so far as it is Q*, and simultaneously to deny that some other individual, which is also Q, is P.

But does not the statement "The king of this kingdom can do no harm" mean the same as "For every x, if x is king of this kingdom, x can do no harm"? Not quite: in the latter statement, it is not asserted that the antecedent is the reason for the consequent, whereas the former does say that *qua king of this kingdom* one can do no harm. This is, moreover, a uniform reason, in the sense mentioned above. Knowing the truth of the latter statement, we have the same reason for asserting that James can do no harm as we have in the case of Charles and George.

Of course, this is not entirely new; the novelty of my proposal consists just in the fact that reference to forms, besides ordinary reference to individuals, can be shown to be a *bona fide* type of reference, in the sense that it allows the setting up of a definition of truth along perfectly Tarskian lines.[26]

Now we are in a better position to appreciate Tait's appeal to arbitrariness in connection with finitism, even if there are still a number of questions waiting to be answered. For instance, can we grasp a form before we grasp any one object having that form? Or can we grasp forms only in the objects – an act which might perhaps be termed "intuition"? And how can we abstract a form from an object? These are all difficult (if well known) questions, and I do not have anything new to say about them. But no matter how one answers them, it is plausible to say that we can grasp a form before we run throught all the objects having that form, perhaps even before we see any one of such objects. This goes a long way towards explaining how we can grasp *in general* how a function operates on any object having a given form, without any need to know which values it yields for which arguments.

The picture of arbitrariness resulting from all this is then the following. What turns an ordinary individual into an arbitrary one, thus justifying the generalization, is our seeing that it possesses a certain form and that it is in virtue of that form that it has the property we are interested in. As soon as we see that form as a reason for having the property in question, we thereby also know that any (other) individual with that form has the same property; as I said above, this is just part of our seeing the form as a reason for having a property. There is no need to grasp any totality of objects in order to understand this, and the finitist can understand it as well as anybody else.

But we can see that an object can possess a given property in virtue of having a certain form, even without examining any object at all. In fact it is not even necessary that we acquire forms only by abstraction from the

things that have them, as Frege clearly saw.[27] This explains why one can infer the universal generalization of a given property, e.g. in a proof by *reductio*, even if there is no actual object having that property.

That we have difficulties in seeing all this might perhaps be due to the formalism we are used to. First order logic seems to be effective in showing that an object satisfies a given property in virtue of some form it has, *by means of the form of some proof* we have found of that fact (e.g., if the proof is such that an application of the rule of Introduction of the universal quantifier can be appended to it). But we do not know how to say this within that same formalism.

NOTES

[1] See K. Fine, *Reasoning With Arbitrary Objects*, Oxford, 1985. M. Santambrogio, "Generic and Intensional Objects", *Synthese* 73 (1987), 637-63. In a way, Montague's theory of quantification can also be considered to be a theory of indefinite objects.

[2] See, e.g., on p. 45 of *The Principles of Mathematics*, Allen and Unwin, London, 1956, the paragraphs about "that curious twofold use which is involved in *human* and *humanity*.

[3] G. Frege, *The Foundations of Arithmetic*, English Translation by J. Austin, Blackwell, Oxford, 1968, pp. 60e-61e.

[4] See "On Concept and Object", in *Translations from the Philosophical Writings of Gottlob Frege*, by P. Geach and M. Black, Blackwell, Oxford, 1970. p. 45.

[5] That horse-hood and whale-hood are objects is, on Frege's view, certain: for him, an object is simply whatever can be referred to by a linguistic expression that behaves as a proper name (in Frege's own extended sense), and there is little doubt that "horse-hood" and "whale-hood" do behave as names. Abstractness, on the other hand, was never considered by Frege as a source of ontological embarrassment.

[6] See, e.g., J. Locke, *Essay Concerning Human Understanding*, Book III, Ch. VIII, Par. 2.

[7] D. Scott, "Identity and Existence in Intuitionistic Logic", in *Applications of Sheaves*, M. Fourman, C. Mulvey, D. Scott, eds. Springer, 753.

[8] "What Is A Function?", in *Translations...*, p. 109.

[9] Fine claims that by the clause "Let x and y be two arbitrary reals" what we really mean is that "[t] here is a unique arbitrary pair of reals, p; it is the independent arbitrary object whose values are all the pairs of reals", p. 19. For a criticism of this solution, see my *Review* of that volume, forthcoming in *Noûs*.

As to Russell, see the following passage: "Thus x is, in some sense, the object denoted by *any term*; yet this can hardly be strictly maintained, for different variables may occur in a proposition, yet the object denoted by *any term*, one would suppose, is unique. This, however, elicits a new point in the theory of denoting, namely that *any term* does not denote, properly speaking, an assemblage of terms, but denotes one term, only not one particular definite term. Thus *any term* may denote different terms in different places. We may say: any term has some relation to any term; and this is quite a different proposition from: any term has some relation to itself. Thus variables have a kind of individuality..." *Principles*, p. 94.

[10] "What Is A Function?", p. 110.
[11] Bas van Fraassen, "Subjective Semantics" in *Versus*, 44-45 (1988), pp. 201-20.
[12] See, e.g., "A Critical Elucidation of Some Points in E. Schroeder's *Vorlesungen ueber die Algebra der Logik*" in *Translations from the Philosophical Writings of G. Frege*, by P. Geach and M. Black, Blackwell, Oxford, 1970.
[13] Jackendoff devotes to this topic a considerable part of his *Semantics and Cognition*, M.I.T. Press, The Bradford Books, 1985. It must be said, however, that not everybody is disturbed by this ambiguity. Russell, for one, noticed that "The word is terribly ambiguous" and distinguished at least five different meanings; "doubtless there are further meanings which have not occurred to me", but he doesn't seem to worry. *See The Principles of Mathematics*, p. 64.
[14] "On Concept and Object", p. 48.
[15] "A Critical Elucidation...", p. 105.
[16] In Dummett's own words: "We understand the univerally quantified statement because we have, as it were, a *general* grasp of the totality which constitutes the domain of quantification – we, as it were, survey it in thought as a whole – and because we know what it is for the predicate to which the quantifier is attached to be true or false of any one arbitrary element of this domain", *Frege, Philosophy of Language*, Duckworth, London, 1973, 517-18.
[17] *Frege*, pp. 517-8.
[18] M. Dummett, *Elements of Intuitionism*, Oxford, 1977, pp. 14-15.
[19] "However, a proof of (x)A(x) does not have to take this simple form. We can have an operation which, applied to any number n, yields a proof of A(n) even though the structure of the proof depends on the value of n. An easy example of this is given by the intuitionistic justification of induction. Suppose that we have a proof of A(0) and a proof of (x) (A(x) —> A(x + 1)), which we may suupose for simplicity to have been obtained by means of a free-variable proof of A(x) —> A(x + 1). Then, for each n, we find a proof of A(n). When n = 1, we apply modus ponens to A(0) and A(0) —> A(1); when n = 2, we first obtain A(1) by the preceding modus ponens step, and then apply modus ponens again to A(1) and A(1) —> A(2); and so on. There is no uniform proof-skeleton (except one which allows explicit appeal to induction); the length of the proof (number of applications of modus ponens) depends on n; but we have an operation which we can recognize as yielding a proof of A(n) for each n". ibid. p. 14.
[20] See K. Fine, op. cit., p. 137 ff.
[21] W. Tait, "Finitism", in *The Journal of Philosophy*, 1981. 524-46.
[22] ibid., p. 528.
[23] ibid., p. 528.
[24] ibid.
[25] On this important topic, see Russell, *The Principles of Mathematics*, p. 91: "If *n* stands for any integer, we cannot say that *n* is 1, nor yet that it is 2, nor yet that it is any other particular number. In fact, *n* just denotes *any* number, and this is something quite distinct from each and all of the numbers. It is not true that 1 is any number, though it is true that whatever holds for any number holds for 1. The variable, in short, requires the indefinable notion of *any* which was explained in Chapter V." And further down on the same page: "There is a certain difficulty about such propositions as 'any number is a number'. [...] if 'any number' be taken to be a definite object, it is plain that it is not identical with 1 or 2 or

3 or any number that may be mentioned. Yet these are all the numbers there are, so that 'any number' cannot be a number at all. The fact is that the concept 'any number' does denote one number, but not a particular one. This is just the distinctive point about *any*, that it denotes a term of a class, but in an impartial distributive manner, with no preference for one term over another. Thus although x is a number, and no one number is x, yet there is no contradiction, so soon as it is recognized that x is not one definite term".

[26] This I showed in my paper "Generic and Intensional Objects:, cit., where formal details are given. It is also shown there that the two notions of reference are not unrelated.

[27] See *The Foundations of Arithmetic*, cit., p. 62^{e}-63^{e}.

PETER SIMONS

LOGICAL ATOMISM AND ITS ONTOLOGICAL REFINEMENT: A DEFENSE

1. INTRODUCTION

Russell employed the term 'logical atomism' for his own philosophy when it was heavily under the influence of Wittgenstein, and it came to be also applied to the *Tractatus*.[1] It is now widely believed that logical atomism is untenable. But the differences between Russell and Wittgenstein show that the term 'logical atomism' is general, not singular. Possibly the objectionable aspects of their theories are not essential to logical atomism as such. I shall attempt to give a form of logical atomism which is both natural and plausible, suggest why it is an attractive view, and take steps in the direction of justifying it against standard objections.

2. LOGICAL ATOMISM, WHAT

The term 'logical atomism' is not exactly transparent. Russell says (RLA, p. 33)[2]

> The reason that I call my doctrine *logical* atomism is because the atoms that I wish to arrive at as the sort of last residue in analysis are logical atoms and not physical atoms. Some of them will be what I call 'particulars' – such things as little patches of colour or sounds, momentary things – and some of them will be predicates or relations and so on.

This is distinct not only from physical atomism, where the atoms are particles, but also psychological atomism of the Humean variety, where they are ideas. The terminology of atoms, compounds, structure, combination, and analysis shows that Russell and Wittgenstein were exploiting the analogy with chemistry. But the analysis is also closely

linked with logic. In fact in the hands of Russell and Wittgenstein logical atomism was a two-tiered system. One tier is concerned with accounting for the truth and falsity of logically complex propositions in terms of logically simpler ones. This part of the theory seems to be mainly the work of Wittgenstein with his account of the truth-functions, although the idea is in principle much older, and can be found for example in Ockham.[3] The second tier is involved with discovering further hidden complexity in apparently simple propositions and continuing the analysis until the ultimate elements are revealed. The model here seems to have been Russell's theory of definite descriptions, which discerned logical complexity in apparently simple propositions e.g. of the form 'The A fs'. Wittgenstein acknowledges this (TLP 4.0031):

> It was Russell who performed the service of showing that the apparent logical form of a proposition need not be its real one.

Though inspired by Russell's theory of descriptions, Wittgenstein's own conception of analysis is different from it and owes more to the chemical analogy.[4] Russell's logical atoms were sense-data and universals. Wittgenstein did not commit himself in the *Tractatus* as to what his logical atoms are. Although I look with favour on the second tier of logical atomism (cf. 10), I shall seriously defend only the first tier.

Logical atomism is a view about the relationship between truths and what make truths true. I am here taking for granted the view that an adequate theory of truth must make use of the notion of that which makes a truth true, or truth-maker.[5] Since the development of truth-theory without truth-makers by Tarski, it has been widely held that entities such as the facts of Russell and Wittgenstein are not needed. But Tarski's view leaves a theoretical lacuna which keeps it from completely explicating truth. Fundamental to Tarski's theory is the concept of a sequence of objects (which need not be individuals) satisfying an open sentence. The most familiar case is where a simple predicate is satisfied by a sequence of individuals. Take the simplest such case: a monadic predicate and an individual satisfying it. Tarski's theory never descends to consider the naive and natural question whether there must be anything in virtue of which this individual satisfies this predicate. For the semantics of formal systems, the question is irrelevant, but if we are interested, as Tarski said he was, in truth, the question needs addressing. Tarski's Convention T is not enough to settle the issue, because it requires only that all T-biconditionals be derivable from the theory of truth in the metalanguage.

But, provided we assure that semantic paradoxes cannot arise, T-biconditionals turn out to tell us nothing about what makes a truth true.[6] The question 'in virtue of what does an individual satisfy a predicate?' arises as soon as one recognizes that the truth of a sentence is the joint outcome of two largely independent factors: on the one hand that about the language which determines what the sentence in question means, and on the other whatever it is in the world which determines that the sentence, meaning as it does, is true or false. So Tarski's theory is in need of supplementation by considerations about the entities in virtue of which propositions are true. The first principle of truth-making is that a true proposition is not in general simply true without further ado, but is true in virtue of something or other, is made true by something or other. That is

TM For all p: p $\leftrightarrow$ for some X, X makes it true that p.

The double arrow here is a co-entailment (necessary relevant coimplication) connective. The variable 'X' ranges, on the nominalistic version of the theory which I prefer, over particulars and arbitrary pluralities of particulars: other versions (such as Russell's) may invoke universals, but I shall not. The notion 'makes it true that' has nothing to do with causality. It has been recently glossed as 'that – exists entails that',[7] which helps to dispel some of its mystique. A truth-maker for p is simply something whose existence entails that p. I shall leave it open whether this gloss is acceptable.

Russell and Wittgenstein called those entities which make truths true 'facts'. I shall follow them in using this convenient word to supplement the more transparent but more barbarous expression 'truth-maker', but I deny that there is a separate ontological category of objects whose *peculiar* function it is to make truths true: truths are made true by sundry particular items from various ontological categories. Facts are just things that make truths true.

A truth is for present purposes a true proposition and a falsehood a false proposition. The reasons for adopting this terminology rather than talking about sentences or statements need a word of explanation. In general, as is well known, sentences as such do not have truth-values, but only as used by particular persons on particular occasions to make statements. The resulting truth-value, whether we ascribe it to the statement (the speech act of uttering this sentence with suitable intent) or derivatively to the sentence itself, is usually the outcome of three factors:

(1) the connection between the sentence itself and its linguistic meaning (the latter being that which a good translation preserves), (2) the circumstances of utterance, and (3) whether things are as they are thereby said to be. For our concerns here only the third factor is important, so we use the term 'proposition' to enable us to abstract from the peculiarities of language and situation of utterance, without being seriously committed to abstract propositions.

The components of propositions which are not themselves propositions I shall call 'concepts'. Their status is similarly merely heuristic. Those concepts which are not complex may be called 'atomic concepts'. For example *white* and *book* might well be atomic concepts, but *white book* is not.[8] Concepts differ in kind according to their semantic role. Particularly important for us are names and predicates (for which the linguistic terminology is appropriated). The semantic value of a name, the contribution it makes to the truth-value of propositions in which it occurs, is given by the individual or individuals it names (its nominata, including the case where it names nothing at all). The semantic value of a predicate is determined by which truth-value it outputs for which individuals (or tuples of individuals) as inputs. Calling names and predicates together *terms*, an *atomic proposition* is one containing only atomic terms. It is the counterpart of a simple predicative sentence, in which no expression explicitly contains or implicitly presupposes (via definition) any propositional connectives, quantifiers, or other logical constants, or any concept which is not a term.

According to TM true atomic propositions have truth-makers. Call the facts which make atomic propositions true 'atomic facts'. What then of true non-atomic propositions? By TM, they too have truth-makers, but TM leaves its open whether we need special facts over and above those we already need for atomic propositions to make non-atomic propositions true. Logical atomism is the view that we do not. To be more precise, we relativize the thesis of logical atomism to a class of propositions.

LA If P is a class of propositions, logical atomism with respect to P is the view that all the true propositions in P are made true by atomic facts.

Wittgenstein, unlike Russell, upheld *universal* logical atomism: for him, *all* true propositions are made true by atomic facts. (Wittgenstein's position is not as strong as it sounds because he had a rather narrow view as to what may count as a proposition.) In what follows I shall not attempt

to make such a universal logical atomism plausible, but to concentrate on non-necessary (contingent) truths.[9] Contingent truths require no other truth-makers than those which make contingent propositions true. There are, with respect to contingent truths, no inherently logically complex facts, no general, higher-order or any other facts except atomic facts.

The form of logical atomism I defend differs in various respects from the different versions of logical atomism advocated by Russell and Wittgenstein. I do not advocate a metaphysics of facts. I am not tying my logical atomism to any views about the nature of a logically perfect or ideal language. The thesis of logical atomism is only tangentially about language, and, if true, would be true whether we had any logical language or not. I reject Russell's empiricist view that the ultimate particulars are sense-date and consider them rather to be physical, though what particulars there are which serve as truth-makers is an empirical question. Unlike Wittgenstein, I do not believe that logical atomism requires us to regard particulars as themselves without parts. The *Tractatus* account of the nature of complex particulars and their relationship to non-atomic facts is a chapter of confusions.[10] Nor do I accept Wittgenstein's view that atomic propositions are logically independent of one another. The thesis of logical atomism is not about definition or truth-conditions or meaning. It is not claimed that propositions containing non-atomic expressions have the same meaning or truth-conditions as or are definable in terms of atomic propositions. The running together of logical, linguistic and ontological analysis was one of the more dire outgrowths of Wittgenstein's besottedness with language.

3. EXAMPLES OF THE AVOIDANCE OF UNNECESSARY FACTS

Suppose that John loves Màry now, on 4 June 1988. Then the same fact or facts which make this true make true the disjunctive proposition that John loves Mary or Mary loves John. (In future instead of 'fact or facts' I shall say just 'facts', allowing the plural to subsume the singular.) If Mary loves John, the facts which make this true will also suffice to make the disjunctive proposition true. There is no need for *disjunctive* facts.

If each loves the other, then the facts which make each separate proposition true suffice *together* to make the conjunction true. There is no need for *conjunctive* facts. This is why in TM we need to include pluralities of facts.

Suppose John does not love Lucy. He need not hate her, and indeed

hating someone is neither necessary nor sufficient for not loving them. He may like her, be indifferent towards her, or never have heard of her. All that we require is that of all his emotive attitudes, none should be love for Lucy. In all the facts E involving John's emotions, none makes it true that he loves Lucy, so all these facts E together make it true that he does not love Lucy. There is no need for *negative* facts. This is highly disputed and I shall come back to it.

Suppose John not only loves Mary but each of her six sisters as well. Then the fact which serve to make each of the six propositions true obtained by replacing the 'x' in 'John loves x' by proper names of each sister in turn, together with the facts making true each of the six propositions similarly obtained by replacing the 'x' in 'x is sister of Mary' by proper names of the six sisters, all cooperate together to make it true that all of Mary's sisters are loved by John. There is no need for *universal* facts to make universal generalizations true. A universal generalization is made true collectively by all the facts which make each of its instances true. This is likewise disputed and will be discussed in greater detail below.

There is no need for special facts making particular generalizations true. If John loves Mary, then whatever facts make this true also that John loves someone, that someone loves Mary, that someone loves someone. Any facts which make an instance of a propositional function true make its particular generalization true. Note that we have spoken of *particular* rather than existential generalizations. In some logical dialects, these two conceptions coincide. In others, they do not. If they do not, we can add that an existential generalization is made true by whatever facts makes any of its instances true, plus whatever existential facts are also required. So if to derive 'there exists x such that Fx' from 'Fa' we require also that a exist, then to make the existential generalization true we need the facts making 'Fa' true and the facts making 'a exists' true. What these latter might be, we shall come to presently.

True identity propositions involving only atomic names do not require their own special truth-makers. There are two plausible stories about the truth-maker for such propositions. Either there is none at all, or the object itself makes the propositions true. In neither case do we need special facts of identity. If we adopt the second alternative, we do indeed have facts (namely the things themselves) which make the identity propositions true, but these facts in no way involve any component corresponding to the identity predicate. We may wish to leave room for both alternatives, having an identity predicate with and another without existential import,

as in certain free logics. One reason for liking the object itself as truth-maker for the self-identity proposition is that a false identity proposition is most plausibly made false conjointly by the two objects named (which thereby make true the difference proposition and the logically equivalent or perhaps identical numerical proposition that they are two objects.)

What should we do about the true singular existence propositions of the form 'a exists'? Here again we have more than one alternative. If we define such propositions, as is usual in those free logics which stay as close as possible to classical logic, to mean 'a is identical with something' then, because of our account of existential and identity propositions, a true singular existence proposition has as truth-maker whatever makes the identity true (either the object a itself, or nothing) plus whatever is additionally needed (if anything) to make the generalization true. However if propositions involving existential generalizations require among their truth-makers not just those facts making their instances true, but also those facts making the corresponding singular existential propositions true, then we have got nowhere, since that was what was to be decided. The simplest solution is such cases is to make the object a itself truth-maker for its own existential proposition. That gives us in this sense singular existential facts, but these in no way involve any component corresponding to the verb 'exists.'

That is my favoured solution, but it does not work for those of Leibnizian or Meinongian bent, for whom not all objects exist. In this case the object a on its own cannot serve as truth-maker for 'a exists', since then all objects would exist. We need some other fact, a special fact of existence. Exactly what form this should take may be left open here. But note that the acceptance of such facts of existence is compatible with the spirit of logical atomism, as long as they are not facts of higher order, as they would be on some Fregean interpretations of singular existence sentences.

4. DISPUTED CASE I: NEGATIVE PROPOSITIONS

True negative propositions, especially true negations of atomic propositions, have always given logical atomists trouble. In *The Philosophy of Logical Atomism* Russell writes (RLA, p. 67),

> One has a certain repugnance to negative facts, the same sort of feeling that makes you wish not to have a fact 'p or q' going about the world. You have a feeling that there are only positive facts, and that negative propositions have somehow or other got to be expressions of positive facts. When I was lecturing on this subject in Harvard. I argued that there were

negative facts, and it nearly produced a riot: the class would not hear of there being negative facts at all.

Russell argues against a paper of Raphael Demos[11] that 'not-p' is to be defined as

for some proposition q, q is true and q is incompatible with p

by pointing out that incompatibility is a relation among propositions, which Russell denies are denizens of reality, and that facts of incompatibility are at least as dubious as negative facts. For these reasons, and in order to provide a *falsehood*-maker for false atomic propositions, Russell prefers to hang onto negative facts. In contrast to many earlier and later philosophers, he regards the issue as a substantive one: his remark from the discussion bears quotation (RLA, p. 71):

It seems to me that the business of metaphysics is to describe the world, and it is in my opinion a real definite question whether in a complete description of the world you would have to mention negative facts or not.

Wittgenstein's logical atomism is more ascetic in this point than Russell's. Admittedly Wittgenstein speaks of negative facts as the non-obtaining of states of affairs (TLP 2.06), which is a timely reminder that Wittgenstein's logical atomism contains an extra layer of ontological structure above and beyond that of Russell, namely the distinction between facts and states of affairs (one which the original translation of 'Sachverhalt' as 'atomic fact' effectively obscured). Wittgenstein himself was prone to forget the subtlety (and hence to endorse the bad translation): in a letter to Russell of 19 August 1919 he writes[12]

"What is the difference between Tatsache and Sachverhalt?" Sachverhalt is, what corresponds to an Elementarsatz if it is true. Tatsache is what corresponds to the logical product of elementary prop[osition]s when this product is true.

This covers only wholly positive (Wittgensteinian) facts, but there are wholly negative and mixed facts as well. But states of affairs (and their corresponding elementary propositions) cannot be negative: in the same letter Wittgenstein retorts with unmistakable (and unfair) derision, "of course no elementary prop[osition]s are negative".[13]

Consider then a false atomic proposition, and its true negation. What makes the former false and the latter true? If we say "The non-existence of any fact that would make the atomic proposition true", we accept facts of non-existence. Maybe all simple negative propositions could be dealt

with by facts of non-existence, but these are still negative, and seem to involve us in the Leibnizian-Meinongian assumption of non-existents. An earlier joint attempt to solve the problem[14] copies Wittgenstein's two-layered approach, introducing special facts of existence and non-existence in addition to other truth-makers, but I find it unconvincing. If we try to modify Demos's incompatibility solution and say that the negative proposition is made true by the existence of a fact whose existence is incompatible with that whose existence would make the atomic proposition true (i.e. express incompability, as Russell intimates, not as a relationship among propositions), then we have a relationship joining an existent with a non-existent fact, which is a queer kind of relationship. It is difficult to escape Russell's conclusion, that incompatibility is either itself negative or an evasion which only complicates without enlightening.[15] Introducing a relation of making *false* alongside that of making true seems not to eliminate the problem, because the most plausible account what it is for X to make p false is that it is for it to make not-p true. So how can we get by without facts of absence or non-existence?

Consider the *monotonicity* of truth-making:[16]

MON If X makes p true, and X is included in Y, then Y makes p true.

Suppose we have a maximally inclusive plurality of truth-makers, a totality of all (atomic) facts. Call it W ('the world': on our view the totality of facts is the same as the totality of things). Then by TM and MON

WM For all p: p $\leftrightarrow$ W makes it true that p.

If we need a truth-maker for a negative proposition, we need look no further than W. If nothing else makes a proposition true, W will.

This solution preserves TM and is much in the spirit of Wittgenstein (cf. TLP 1.11-2, 2.05). What is true and false is determined by the facts, and by their being all the facts: whatever is not made true (by anything) is thereby false. Their being all the facts is not a further fact, as argued in the next section. The notion of the totality of all facts (things) may raise some doubts, though since the things in question are in my view all physical, I do not see that the notion engenders paradox. This solution is important for those interested in maintaining TM, and it is my favorite for that reason, but logical atomism can perhaps get by with less.

A weakening of TM has been suggested by Bigelow:[17]

If something is true, then it would not be possible for it to be false unless certain things were to exist which don't, or else certain things had not existed which do.

This maintains the view that truth is supervenient on being, but it lacks the metaphysical bite of TM, which can be used against certain standard arguments for universals.[18] So on balance I prefer to support TM in all its glory.

5. DISPUTED CASE II: UNIVERSAL GENERALIZATION

If neither particular generalizations nor negative propositions require their own facts, *a fortiori* neither do universal generalizations. But let us confront the case in its own right. The obvious candidate for the truth-makers for a universal generalization of the form $\forall x Fx$ are the truth-makers for the instances, Fa, Fb, and so on. Russell says of this solution however (RLA, pp. 92-4):

You can never arrive at a general fact by inference from particular facts, however numerous. [...] It is perfectly clear, I think, that when you have enumerated all the atomic facts in the world, it is a further fact about the world that those are all the atomic facts there are about the world [...]
The same thing applies to 'All men are mortal'. When you have taken all the particular men that there are, and found each one of them severally to be mortal, it is definitely a new fact that all men are mortal [...] it could not be inferred from the mortality of the several men that there are in the world.

Russell's argument is based on the mistaken assumption that if the facts $F_1,\ldots,F_n$ make propositions $p_1,\ldots,p_n$ true, then they cannot together make the proposition q true unless q follows logically from the conjunction of the p_i. I suspect one reason for Russell's view here is his vacillation in the use of his term of art 'fact'. Some occurrences of this seem to mean what we expect it to mean, others however are closer to the usage recommended by Frege and Ramsey, as meaning 'true proposition'. It is notable that the beginning of the fragmented quotation given above comes when Russell is discussing whether general *propositions* follow from particular (i.e. singular) ones, and later on the same page he says (RLA, p. 92),

You never can arrive at a general proposition by inference from particular propositions alone. You will always have to have at least one general proposition in your premisses.

Only then does he say (RLA, p. 93),

I come now to a question which concerns logic more nearly, namely, the reasons for supposing that there are general facts as well as general propositions.

Whatever the reasons for Russell's view, I think it is wrong. Contrast what Wittgenstein has to say on the same issue. I give a mosaic of quotations, which could be extended:

4.41 Truth-possibilities of elementary propositions are the conditions of the truth and falsity of propositions.

4.51 Suppose that I am given *all* elementary propositions: then I can simply ask what propositions I can construct out of them. And there I have *all* propositions, and that fixes their limits.

4.52 Propositions comprise all that follow from the totality of all elementary propositions (and, of course, from its being the *totality* of all them all). (Thus, in a certain sense, it could be said that all propositions were generalizations of elementary propositions.)

5.5265 The truth or falsity of *every* proposition does make some alteration in the general construction of the world. [...]

6.1231 The mark of a logical proposition is *not* general validity. To be general means no more than to be accidentally valid for all things. [...]

6.1232 The general validity of logic might be called essential, in contrast with the accidental general validity of such propositions as 'All men are mortal'.

The point at which Wittgenstein comes closest to saying what Russell says is 4.52. But while Wittgenstein *requires* that if a proposition $\forall x Fx$ is to be true, all its instances must be true, he does not require an additional premiss "There are no objects other than a, b, c, ... etc." Indeed, since on Wittgenstein's view such a thing cannot be said at all, *a fortiori* it cannot be a premiss of the argument inferring $\forall x Fx$ from the complete collection of its instances. That this is not what is normally meant by following logically is marked by Wittgenstein's "in a certain sense", since of course the addition of an extra individual to the domain of quantification could change the truth-value of the universal proposition. That this is a correct interpretation of Wittgenstein is confirmed by the other quotations. Note in particular that generality, i.e., universality, *is* no more than truth of all

instances, what Wittgenstein calls 'accidental general validity.'[19]

So Wittgenstein, unlike Russell, did not consider there to be any need for universal facts over and above atomic facts. And I think he is right.

6. OTHER HIGHER ORDER FUNCTORS

The propositions looked at so far are familiar from first-order predicate logic and its variants. If we are right about these propositions, then any proposition expressible in terms of first-order predicate logic with identity is likewise made true only by atomic facts. This includes propositions involving numerical quantifiers. But there are functors of higher order which are not definable in first-order terms. The monadic quantifer 'there are finitely many' and the dyadic majoritative quantifier 'most' are examples. Is there a barrier here which the atomic facts cannot overcome? I think not.

Note first that the question of definability has to do with valid mutual implications, having the same truth-value in all models. We saw in connection with negation and universal generalization that questions of what makes a proposition true and questions of what logically implies or is implied by the proposition come apart. So although 'there are finitely many As' may not be *defined* in first-order terms, we may always express in first-order terms the propositions whose truth in the sense mentioned suffices for the truth of the finiteness proposition, assuming it is true. If there happen to be three As, call them X, Y and Z, then the following seven propositions are true:

X is an A
Y is an A
Z is an A
X $\neq$ Y
Y $\neq$ Z
Z $\neq$ X
For all w: w is not an A or w = X or w = Y or w = Z

Since 'There are finitely many As' follows from this, whatever facts make the conjunction of these seven propositions true makes the finiteness proposition true. And we have argued that propositions of all three types represented above require only atomic facts, so the finiteness proposition in this case requires only atomic facts. But all true finiteness propositions follow from some true definite numerical proposition, so no true

finiteness proposition requires other than atomic facts.

Similarly, a proposition of the form 'Most As are Bs' is true iff there are more As which are Bs than As which are not Bs. Suppose there are five As, three of which are Bs. Suppose these As are called V, W, X, Y and Z, and that X, Y and Z are the Bs. Then the following are true:

V is an A	
W is an A	
X is an A	X is a B
Y is an A	Y is a B
Z is an A	Z is a B

Each of these propositions has one or more facts which make it true. These do not suffice on their own to make it true that most As are Bs, since their being true is compatible with there being another seven As, not listed here, none of which is a B. So for the truth of 'Most As are Bs' to be guaranteed by these, we need that nothing further is an A. We also need to add that X, Y and Z are three, and not two or one. The case is then here not different in principle from that for the finiteness quantifier. At least we cover all cases where there are only finitely many As.

Suppose there are infinitely many As. What makes this true? Answer: the totality of facts making true each true proposition of the form 'X is an A and Y is an A and X $\neq$ Y'. That there must be infinitely many such facts is of course not surprising, but is no barrier to the logical atomist reduction. The distinction between finite and infinite collections of facts is logically insignificant, something which, incidentally, should make us suspicious of attempts to drive a wedge between first-order and higher-order logic in terms of their philosophical acceptability.

Consider again the majoritative quantifier. What if both the As and the Bs are infinite in number? For our general solution to work, we should need to consider in at least one case a non-denumerable infinity of facts. The standard second-order definition of this quantifier, in terms of there being an injection to the ABs from the A non-Bs but not vice versa, is very short and easily understood. By contrast, if we persist with atomic facts, we have an unsurveyable multitude. But the transition from denumerable to non-denumerable collections is not significant. We are concerned not with expressing propositions corresponding to each of these facts, analyzing their meaning, or expressing their truth-conditions in terms of atomic propositions. Our inability to spell out in detail what the truth-makers of such propositions are shows clearly why we need forms of

expression going beyond atomic propositions and their molecular combinations. The finite case required only atomic facts, and the transition to cases of larger cardinality makes no difference of principle. If infinite domains do not deter us for the quantifiers binding individual variables, there is no reason why they should deter us for higher-order quantifiers.[20]

7. STATISTICAL GENERALIZATIONS AND PROBABILITY

It is probably obvious by now what I incline to say about statistical generalizations and probability propositions. Take the former first, and consider again our example of five As, three of which are Bs. The atomic facts already mentioned suffice to make true the empirical statistical generalization that 60% of all As are Bs. Statistical propositions are made true by the atomic facts concerning members of the population according to proportion of favorable "outcomes" to all "outcomes".

Probability propositions give more trouble. Firstly I want to bracket questions about strength of belief. I think we may legitimately ask about such things, but I think it is a chapter of the theory of truth-makers for mental attitudes and is slightly off our track here, where I am concerned with objective probability. The most attractive theory of probability propositions for a logical atomist is not doubt a form of relative frequency theory, such as is found in Venn and von Mises, and whose most complete defence is probably that of Reichenbach. On this view, probability propositions are interpreted as in fact statistical propositions, usually about 'long runs' of like events. For a "short run", take the case where of the five As, three are Bs. If you pick an A "at random", as they say, the probability is 0.6 that it will be a B. The connection between the second sentence and the first is logical: they are logically equivalent and the expression which does most to bridge the gap is 'at random'. Since statistical propositions present the logical atomist with no problem, the argument goes, neither do probability propositions. But is the relative frequency view correct? There are a number of arguments against, though I agree with van Fraassen[21] that they are not as strong as has often been thought.

A standard argument against the relative frequency theory is that it makes it meaningless to assign probabilities to single events. It is not that there is no way of assigning a probability to the event, but, because a single event may be subsumed under many different descriptions, each of

which determines a different reference class, the ratio of favorable cases to the total will vary with the description. Only I cannot see why this is meant to undermine the relative frequency view. On the contrary, the view makes it appear incorrect to ever assign a probability to a single event *per se*.

More worryingly, a number of writers have shown that Reichenbach's frequency theory does not square fully with the axioms of probability theory: in particular, the domain of definition of long run relative frequencies is not countably additive.[22] Van Fraassen goes on to offer a modified frequency interpretation which makes use of a number of modal idealizations to square relative frequencies with Kolmogorov's axioms. Van Fraassen's idealizations are modal in character, having to do with the relative frequency of successful outcomes in infinitely long runs of ideal experiments. The modal aspect effectively puts this out of reach of our version of logical atomism. But van Fraassen never questions Kolmogorov's axioms, because they are what everyone, especially physicists, use. Most writers on the philosophy of probability see them as the one bright, stable spot in a sea of conflicting intuitions and interpretations. But why should they be sacrosanct? Perhaps, like real and complex analysis, mathematical probability theory as applied to empirical matters embodies heuristically useful but strictly false simplifications and idealizations. At any rate, I do not think that the fact that Kolmogorov's axioms do not perfectly fit the relative frequency view of probability necessarily scuppers logical atomism. Rather it shows that if we wish to hang on to the truth of probability propositions, to the extent that they do not square with empirical statistics, they are not contingent at all, but necessarily true, being partly constitutive of the concept of probability.

Probability propositions are a hard nut for logical atomism to crack, and perhaps one ought to admit at least some of them among the atomic propositions.

8. LAWS OF NATURE AND CAUSALITY

A rather different set of considerations is raised by the question of laws of nature, which I understand as certain kinds of true universal or statistical propositions, fulfilling further conditions which are not relevant here. Laws being true, they are to be distinguished from those generalizations of such a form currently believed by a suitably weighted sum of the scientific community to be true or approximately true, or which they find

useful in their scientific theorizing, and which are therefore often called 'laws of nature'. Most of these are false, and are therefore not laws on my understanding.[23]

If laws of nature are among true universal or statistical generalizations of the form outlined, then providing we can give a logical atomist account of statistical generalizations, laws of nature are made true by the facts verifying their instances, just as in the case of other universal generalizations. However it is often claimed, and is indeed perhaps the majority view, that laws of nature are more than just accidentally true, that while not logically true, they nevertheless have a modal status which lifts them above accidentally true generalizations. In other words, there is a genuine modal difference between a true generalization of the form 'For all x: ---x---' and a true law of nature of the form 'Of natural necessity: for all x: ---x---'. The latter can be false although the former are true. To take an example from Meinong: 'Nothing is an ivory sphere 10 meters in diameter' is a true purely descriptive empirical generalization, but is, according to the majority view, not a law of nature, so its natural-modal strengthening is false. By contrast, it is a law of nature that no body with non-zero mass is accelerated through the speed of light. Not just that it never occurs: it cannot (naturally, not logically) occur.

We may admit that the proposition about the speed of bodies is not logically true. But are we prepared to go as far as Wittgenstein, who put the matter in a nutshell: "There is only *logical* necessity" (TLP 6.37)? Here again I now think the answer is 'Yes'. We should accept that natural necessity is, as Russell and Wittgenstein both think, superstition, nothing but the ghost of a departed celestial monarchy. This view has been ably defended by Swartz,[24] and I refer to his discussion. The arguments that not all true purely descriptive statistical generalizations are true statistical laws may be met by similar considerations to those meeting the claim that not all true purely descriptive universal generalizations are laws of nature. Just as there is no natural necessity, so too there is no natural stochastic tendency or propensity as Popper has claimed. I take the relative unpopularity of Popper's theory as indicating that natural propensity is more keenly felt to be occult than natural necessity.

If this is right, we need no other truth-makers for laws of nature than we need for true universal and statistical propositions, which means logical atomism does not founder on laws of nature.

One thing not much discussed by Swartz is causality. It would be in line with the regularity view of laws of nature to hold that there are no causal

laws above and beyond true Humean constant conjunctions. Causal necessity too, as a subspecies of natural necessity, may go the way of other occult forces. But what of causality itself? Surely when we say that E caused F we are imputing more than that E and F belong to species S and T of such kinds that every S is suitably followed by some T. By suitable choice of species descriptions we could always manufacture spurious causal connections because we can so describe E that the species term it falls under is fulfilled by E alone. But not every pair of successive events is such that the first causes the second. What should we do?

One possible solution which does not require natural necessity or any occult power, yet which conforms closely to what makes us distinguish causal from coincidental successions is to say that in causal successions there is a transfer of energy from the causal antecedent to the causal consequent, in a non-causal succession there is not.[25] Such transfer is to be understood literally. Some of the kinetic energy of one particle is transferred to another with which it collides, some of the electrical energy passing through a bulb filament is transformed into light photons, a battery's chemical energy is transformed in part into electrical energy and this in turn in part into the kinetic energy of a motor driving a vehicle. By contrast, when I pray to the Goddess Fortuna to let me win this week's lotto, the energy expended in my synapses neither reaches the Goddess (because she doesn't exist) nor does any of it start a chain of transfers which eventually influences the way the balls fall. Indeed superstitions of this sort may be viewed as indirect tributes to the energy-transfer theory: "Since my sphere of direct influence is restricted, maybe a suitably structured event will be picked up by the ultrasensitive antennae of the Gods, and amplified and modulated so as to influence the fall of lotto balls." The energy-transfer account needs filling out, but it strikes me as initially plausible and also as explaining, without natural necessity, much that the Humean theory of causation cannot. How then do we explain the fact that energy transfer is so effective? The short answer is, because that's what we take effectiveness to consist in. No more than anything else does energy *compel* and *force* or *make* something happen. When energy is suitably transferred (and here we have to fence the account round with the *ceteris paribus* clauses which bedevil all natural necessity theories of causation), things of the right kind just invariably do happen (and it's because in general they don't invariably happen that we need the *ceteris paribus* clauses!)

9. APPLIED MATHEMATICS, DISPOSITIONS, AND OTHERS

Propositions which describe natural phenomena with the help of mathematical tools such as real and complex analysis, function theory and the like are of course indisputably important for natural science. However the way in which standard textbooks present such propositions (usually in the form of equations) obscures their semantic status almost completely. An ascription of a particular physical magnitude to an object (e.g. the mass of a body, the electric charge on a conductor) is in principle an atomic predication with a special kind of predicate, one which belongs to a family which constitute what Frege called *a magnitude field*.[26] Other quasi-mathematical statements, e.g. differential equations describing the variation or covariation of magnitudes are usually presented as though they were bits of pure analysis, the nature of the magnitudes involved being something to be gathered from the accompanying prose. Equations intended to state laws are implicitly universal quantifications, their appearance notwithstanding.[27] So if individual propositions about magnitude are amenable to atomistic treatment, so are the universal quantifications for which these provide the ultimate instances. In principle, I see no difficulty there. Like most putative laws of nature, they will not be strictly true. And interlarded with them we have statements of pure mathematics, which serve to transform equations, give the form of solutions and the like. They are, if true, then necessarily true, and I am shirking the question what (if anything) make *them* true.

Another class of propositions concern dispositions. To the extent that these are not modal, I suggest their truth-makers are to be found among the underlying states of things which form the non-modal basis for dispositional propositions. To mention only the standard examples: what makes it true that a certain object is brittle are facts concerning the object's microscopic material structure, what makes it true that John is irascible are facts about his overall mental and emotional state, which in suitable circumstances contribute causally to his outbreaks of anger. Because meaning and truth-making are skew to one another, there is no need to strain here for any kind of synonymy or semantic reduction. Dispositional propositions and propositions about the dispositional bases have quite different meanings, but it so happens that the same things make both true.[28]

I take it the kinds of non-atomic proposition covered to date are among the most important. There may be other sorts that I have missed.

Whatever I have left out (and I know of no satisfactory enumeration of the forms of contingent propositions) I must rest with the tentative inductive conclusion: so far, so (say I) good. The atomistic hypothesis may be put in the form of a challenge: give me a putatively true proposition containing a higher-order functor and I shall endeavour to give an outline description of how it can be made true by atomic facts. If I cannot meet the challenge, I will give in.

10. RESOLUTION AND ULTIMATE FACTS

So far I have not done any delving inside propositions which are not obviously logically complex or modal or higher-order. It was a feature of Russell's and Wittgenstein's logical atomisms that they went further and offered analyses or reductions which banished or replaced common or garden individuals in favour of simpler but more esoteric ones – sense-data in Russell's case, objects-we-know-not-what in that of Wittgenstein. Both versions have come in for plenty of criticism. This second tier is an optional extra for logical atomism as here defined. One can do without it. Nevertheless, are there are justifications for such forms of analysis? And if so, what would an acceptable form look like?

The important notion of simplicity in this case is no longer conceptual but ontological, as logical atomists and near-relatives such as Leibniz have always maintained. I shall work with an example. Consider a simple present-tense monadic predication about an individual, 'John has influenza'. What makes this true is in fact rather a complex matter, involving as it does the presence of large numbers of a certain kind of virus in John's respiratory tract and the effects their presence has wrought in him, the production of antibodies by his immune system, the bodily reactions which include some or all of the usual symptoms of influenza, fever, chills, laryngitis, aches, weakness, and depression. The uncovering of the complexity of facts of the kind of those making it true that John has influenza is not the business of semantics but of medical science.

Of the facts required to make it true that John has influenza, some concern the presence of a certain kind of virus. What makes it true that John has influenza includes as proper parts many facts of the sort which make it true that this or that body in his respiratory tract is an influenza virus. There are many strains of influenza virus, so what makes it true that individual B is a member of one such strain is again a complex of facts involving the structure of the virus and the structure of its DNA or RNA molecule.

We can see than that considerations of truth-makers for propositions about states, events and processes involving large and complex individuals naturally leads, as Russell and Wittgenstein thought it did, to considerations of truth-makers involving parts of these individuals. This is the procedure we find in highly schematized form in Wittgenstein's *Tractatus*, albeit wrongly mixed up with questions of linguistic analysis.

Suppose we had the capacity to trace and describe to any desired degree of attainable accuracy what it is that makes a given empirical proposition true. Call this idealized process *resolution*.[29] Because the proposition we start out with will usually be one expressed in a natural language, it will usually be vague in its truth-conditions, but suppose it is nevertheless unequivocally true. In transferring our attention from the proposition itself to specifying what actually makes it true we do not preserve truth-conditions at all, so the vagueness of the starting proposition itself offers no reason for supposing there is the same sort of vagueness attaching to the facts themselves.

In the case of John's influenza, we shall have a very long conjunction of propositions detailing the position, nature, changes and effects of millions of individual flu viruses. Though not every fact involving every individual concerned is relevant to the resolution, there will still be an astronomical number of facts involved. Could the process of resolution go on forever? Why must an analysis involve delving into details at best marginal to our capacity to describe something as 'influenza' and to our practical interest in John's well-being? Because the question "What makes that true?" is not fully answered as long as in giving an account we need to bring in other propositions for which we do not yet have the answer to what makes *them* true. Like the early Wittgenstein, I am persuaded the process theoretically terminates, though like him I do not see how to argue for this thesis in any non-circular way. What then are the general characteristics of the ultimate facts resolution would uncover?

Some of the atomic propositions involved in any resolution will be purely classificatory, saying what a certain object is. Many propositions of this form are not ultimate. That B is an influenza A2 virus is made true by a package of more fundamental facts concerning the parts and moments of B, their nature and relationships to one another. Suppose then we have an individual C and a purely classificatory proposition 'C is a d'. What makes this true if the proposition is not further resolvable into propositions about C's parts and moments? The most obvious candidate is surely C itself. Another kind of proposition likely to be found in any

completed resolution is a statement of magnitude, such as that C has a mass of 10.65 gm., of the sort mentioned in the last section. What makes such propositions true? The aspects or moments of an electron which make it true that it has such and such an electric charge have a position in a certain magnitude range (independent of the system of classification and measurement of this range) but that this moment has this value has no further truth-ground–it just has it. The individual moment is as it is just because it exists. Probably only certain moments (non-independent particulars, tropes) have such an ontologically "thin" structure that they are as they are purely in virtue of existing.

There are ultimate individual relations (relational moments) as well as ultimate individual properties. Resemblance relationships such as obtain between moments at different points of a single magnitude scale (such as different electric charges) are candidate examples. It is not easy to give a good example of an ultimate relationship. In general ultimate facts need more thought, and I mention them here only to indicate how things might develop and to indicate how Russell and Wittgenstein may be retrospectively supported.

11. CONCLUDING REMARKS

In many important points upon which Russell and Wittgenstein differed I find myself closer to Wittgenstein. Russell's acceptance of universal, existential and negative facts make him a very feeble-hearted advocate of logical atomism: about all that he is clearly rejects are disjunctive facts. He is also openly Platonist, whereas Wittgenstein in the *Tractatus* was either neutral or, as I incline to read him, nominalist. On the other hand Wittgenstein's prohibition of metasemiotic language is too restrictive; there are facts about the use of signs and the relationship of signifying which can make metasemiotic propositions true. Their nature is something which I have left aside, as I have also left aside questions about mental attitudes,[30] analytic propositions and the propositions of logic and mathematics, all of which are a challenge to logically atomistic treatment. The challenge is worth taking up if logical atomism proves itself to be defensible among more straightforward empirical propositions. But while I think the arguments for logical atomism are better than is usually admitted, I remain to be more securely convinced, and my defense is to be taken in that experimental spirit.

NOTES

[1] Russell's first use of the term 'logical atomism' probably predates the influences of Wittgenstein: cf. Russell 1911, p. 56.
[2] References to Russell's *Logical Atomism* (RLA) and to *Tractatus Logico-Philosophicus* (TLP) are given in the text. Other references are given in footnotes.
[3] Cf. Freddoso, 1980.
[4] Cf. Griffin 1964, pp. 41-61, Simons 1985.
[5] Cf. Mulligan, Simons and Smith 1984, also Bigelow 1988, pp. 121 ff., Fox 1988.
[6] Cf. Etchemendy 1988, pp. 56 ff.
[7] Cf. Bigelow 1988, p. 125, Fox 1988, p. 189.
[8] I admit there is a problem in deciding what is simple. The convention is adopted of using Italics to name the concept the italicized word expresses. Note that by definition, propositions contain no indexical components.
[9] Since I consider S5 to be the modal logic which captures our alethic modal conceptions this means that all propositions of the form 'Possibly p' are also excluded from consideration.
[10] Demos 1917.
[11] Cf. Simons 1985.
[12] Wittgenstein 1974, p. 72.
[13] Ibid., p. 73.
[14] Mulligan, Simons and Smith 1984, pp. 316-8.
[15] I speak with some pain, as a former adherent of the incompatibility view.
[16] Cf. Mulligan, Simons and Smith 1984, p. 315, formula (10).
[17] Bigelow 1988, p. 133.
[18] Bigelow 1988, pp. 135 ff., Fox 1988, pp. 195 ff.
[19] This argument is borrowed from Swartz 1985, pp. 84-8.
[20] The application of infinitistic considerations to empirical propositions is in any case limited, since those putative truths asserting the existence of large infinite collections come from mathematics, and if interpreted to apply to empirical objects are mainly false.
[21] Cf. van Fraassen 1980, p. 181.
[22] Ibid., P. 184.
[23] Cf. Swartz 1985, pp. 3 ff. and Cartwright 1983 on why most so-called laws are false.
[24] Swartz 1985.
[25] Cf. Fair 1979.
[26] Cf. Frege 1903, p. 158, Simons 1987, pp. 31-3.
[27] Cf. Carnap 1958, p. 169.
[28] Cf. Mulligan 1991, esp. §8
[29] Cf. Simons 1985, pp. 215-7.
[30] On truth-makers for attitude sentences, cf. Simons 1989.

REFERENCES

Bigelow, J. 1988. *The Reality of Numbers*. Oxford: Clarendon Press.
Carnap, R. 1958. *Introduction to Symbolic Logic and Its Applications*. New York: Dover.
Cartwright, N. 1983. *How the Laws of Physics Lie.* Oxford: Clarendon Press.

Demos, R. 1917. "A Discussion of a Certain Type of Negative Proposition", *Mind* n.s. 26, 188-96.

Etchemendy, J. 1988. "Tarski on Truth and Logical Consequence", *Journal of Symbolic Logic* 53, 51-79.

Fair, D.F. 1979. "Causation and the Flow of Energy", *Erkenntnis* 14, 219-50.

Fox, J.F. 1987. "Truthmaker", *Australasian Journal of Philosophy* 65, 188-207.

Freddoso, A.J. 1980. "Ockham's Theory of Truth Conditions". Introduction to *Ockham's Theory of Propositions. Part II of the Summa Logicae.* Notre Dame: Notre Dame University Press, pp. 1-76.

Frege, G. 1903. *Grundgesetze der Arithetik II.* Jena: Pohle. Reprinted Darmstadt: Wissenschaftliche Buchgesellschaft, 1962.

Griffin, J. 1964. *Wittgenstein's Logical Atomism.* Oxford: Clarendon Press.

Mulligan, K., 1991. "Dispositions: Their Bases and Correlates", in: P. M. Simons, ed. *Essays on Meinong.* Munich: Philosophia.

Mulligan, K., Simons, P.M. and Smith B. 1984. "Truth-Makers", *Philosophy and Phenomenological Research* 44, 278-321.

Russell, B. 1911. "Le réalisme analytique". *Bulletin de la Société Française de Philosophie* 11, 53-83.

______ 1972 (RLA) *Russell's Logical Atomism*, ed. D. Pears. London: Fontana/Collins.

Simons, P.M. 1985. "The Old Problem of Complex and Fact". Theoria 5, 205-26.

______ 1987. "Frege's Theory of Real Numbers". *History and Philosophy of Logic* 8, 25-44.

______ 1989. "Wahrmacher fur Aussagen uber propositionale Einstellungen", in: G. Falkenberg, ed., *Epistemische Ausdrücke und propositionale Einstellungen.* Hamburg: Meiner, 199-248.

Swartz, N. 1985. *The Concept of Physical Law.* Cambridge: Cambridge University Press.

van Fraassen, B.C. 1980. *The Scientific Image.* Oxford: Clarendon Press.

Wittgenstein, L. 1961. (TLP) *Tractatus Logico-Philosophicus.* London: Routledge & Kegan Paul.

______ 1974. *Letters to Russell Keynes and Moore.* Oxford: Blackwell.

INGVAR JOHANSSON

INTENTIONALITY AND TENDENCY: HOW TO MAKE ARISTOTLE UP-TO-DATE

1. INTRODUCTION

There is no doubt much in Aristotle's world-view that is far away from the truth. His astronomical theory, which places the Earth in the center of the universe and regards the superlunar world as unchangable, is irreversibly gone. The same goes for his non-evolutionary biology. I hope the same fate is reserved for his views on women and natural slaves. In ontology proper, however, I think it is quite the other way round. Here, we need more Aristotelian ideas.

In my opinion, Aristotle's views on universals come fairly close to the truth. Today, looking at the spread of post-structuralist and post-positivist denials of a language-independent, structured world, I find an Aristotelian immanent realism important not only within philosophy, but in a broader cultural context as well. The view that universals can exist *in re* will not, however, be argued for in this paper. It will simply be presupposed.

My discussion will concentrate on another typical Aristotelian idea, that of natural purpose or intrinsic self-change. This does not mean that my remarks will be exegetical or historical. The primary driving force behind this paper is not my reverence for Aristotle, but a problem which a simultaneous interest in the philosophy of physics and phenomenological philosophy has made me aware of: the problem of distinguishing between the categories of tendency and intentionality. Strictly speaking, this problem belongs primarily to material ontology, but its solution requires some formal ontology at the same time as it indicates a lacuna in formal ontology.

2. THE PROBLEM

During the last decades, the philosophy of physics has witnessed a revival of concepts like tendency, capacity, power and propensity.[1] Sometimes these concepts, especially that of tendency, are meant to refer to entities which have a kind of directedness and which undergo self-produced changes. I shall argue, firstly, that such an Aristotelian concept of tendency cannot possibly be dispensed with, and, secondly, that it makes a customary characterization of intentionality problematic. A tendency has *directedness*, but directedness is often used as a *differentia specifica* of intentionality. The true graphical representation of both tendency and intentionality is the arrow. Therefore, if tendency and intentionality are different categories, neither of them can be characterized merely by the concept of directedness.

Phenomenological philosphy has, I think, for two reasons neglected the problem of how to distinguish between tendencies and intentional acts. First, since most phenomenological philosophers have, as they say, 'bracketed' the natural sciences, they simply cannot see the problem. Second, if they had paid attention to physics, they would probably have met the widespread but false opinion that physics, from Newton and onwards, has been freed from the category of tendency. Most philosphers of physics, on the other hand, have not grappled with the problem because they have confined their interests to the natural sciences. This is not, however, true of Bhaskar, Harré and Popper, but they have for some other reason not perceived the problem; probably because of too vague a conception of intentionality.

As far as I know, there is only one pholosopher who has found the similarity between tendencies and intentional acts (and states) interesting. That is David Armstroing.[2] He employs the similarity between them, i.e. their directedness, in an attempt to reduce intentionality to tendency. According to Armstrong, intentional states are no different in kind from physical states. His argument, very briefly, is that since causality can involve tendencies, i.e. directedness, causality can also explain the directedness of intentionality. And so, he maintains, the causal theory of mind is not threatened by the existence of intentional states.

I am opposed to Armstrong's reductionist position, but I think his argument shows the need for a more detailed and accurate delineation of intentionality. Merely to talk of directedness is not enough.

3. ARISTOTLE

Before discussing tendency vis-à-vis intentionality, I shall say a few words about Aristotle's conception of directed or purposeful intrinsic self-change. This conception is to be understood against the background of three distinctions: (a) between artificial and natural change, (b) between efficient and final causality, and (c) the distinction between actuality and potentiality.

(a) According to Aristotle, everything has a nature, i.e. something which makes the thing *what* it essentially is. When a thing undergoes a change this change may be caused by its nature, but not necessarily so. A change may also be caused by something external to the thing's nature, usually another thing. In such a case there is artificial change; in the former case there is natural change. The distinction between artificial and natural changes, it should be noted, is not identical with the distinction between changes which are caused by something *spatially* external and *spatially* internal, respectively. Something which is spatially internal may none the less be external with regard to the thing's nature. Sickness due to a virus which has entered our body is an artificial change. Growing older is a natural change.

(b) When a sculptor hews out a statue according to Aristotle, the final cause is the idea the sculptor has of the statue. The efficient cause is made up of his hands and tools. The final cause is here external to the thing which is being changed. When a child grows up, or when an acorn becomes an oak, however, the final cause is internal. But independently of whether it is internal or external, the final cause is goal-directed. Final causality always involves directedness.

When the final cause is internal it can in and of itself generate changes, but when it is external it has to be mediated by an efficient cause. If we are to believe Aristotle, the final cause of the acorn changes in and of itself the acorn into an oak, but the sculptor's idea cannot in and of itself change the marble block into a statue. There has to be an efficient cause in between the final cause in the sculptor and the change in the marble block. The final cause, the sculptor's idea, causes directly his hands to move but only indirectly the coming into being of the statue.

If one takes into account the fact that *external* final causality is mediated by efficient causality, the distinction between natural and artifical change runs parallel with that of final and efficient causality. This means that natural changes are always due to final causality, and that

artificial changes are always due to efficient causality. The concept of purposeful intrinsic self-change is then extensionally equivalent both to the concept of natural change and to that of final causality.

(c) The Aristotelian distinction between actuality and potentiality involves in fact a trichotomy. We ought to distinguish between actuality, potentiality and potency. An acorn has, first, a lot of actual properties like weight, shape and colour. Second, it has a *de re* possibility (potentiality) to become an oak. But, third, over and above this potentiality, it also has a tendency or potency to become an oak. The acorn 'strives' to change itself into an oak.

To my mind, the distinction between potentiality and potency/tendency makes visible a conflation in the Aristotelian concept of final causes. A distinction has to be made between a final cause in itself and the goal towards which the final cause is directed. The goal is a *potential* property, but the final cause itself, the potency or tendency, is an *actual* part of the thing in question. These properties are related in such a way that the potency/tendency generates changes only as long as its goal exists only potentially. When the goal is actualised, the potency/tendency passes out of being.

There are two fundamentally different kinds of actual properties: ordinary ones and tendencies. Tendencies are actualities not potentialities. Even when one tendency counteracts another tendency in such a way that no change occurs, both these tendencies are actual. A tendency is identical neither with the changes (of ordinary actual properties) it tends to produce nor with the (potentially existing) goal towards which it tends.

Tendencies can, like ordinary properties, endure, which means that a tendency may be regarded as a state of a substance. When the acorn grows into the oak, it changes its ordinary actual properties most of the time, but the tendency to grow remains the same. The tendency is non-changing. In one sense, therefore, the acorn can be said to be in a state of growth, i.e. in a *state of change*. In ontologies where enduring tendencies are allowed, the concept 'state of change' has a non-contradictory interpretation.

4. NEWTONIAN SELF-CHANGE

What happened to final causality in the post-medieval anti-Aristotelian revolution? The concept had to leave science, the saying goes. Not all sayings, however, are true.

In Newtonian corpuscularism, as in all kinds of atomism, each corpuscle or atom is an indivisible and unchangeable unit. By definition, they cannot change either qualitatively or quantitatively. Atomistic ontologies contain only one kind of change, change of place.

According to Newton's first law of motion, a moving body not affected by any forces continues of itself to move along a straight line with constant speed. In physics, such uniform motion is called inertial motion. Trivially, inertial motion involves changes, changes of place. Non-trivial, however, is the fact that it involves self-changes. In the absence of forces a body in motion will *in and of itself* continue to change place. Right in the middle of Newtonian mechanics, the Aristotelian notion of self-change has survived.[3] Today, Newtonian mechanics has been superseded by relativity theory and quantum mechanics, but the concept of inertial movement has not disappeared. It plays a prominent role in relativity theory. The fact is that not only Aristotelian physics, but classical and modern physics as well, presupposes the concept of self-change or self-movement. If physics is taken realistically there are in the world movements *causa sui*. We have an 'argument from physics' in favour of *causa sui*.

The reason why inertial motion has not been properly conceptualised as *causa sui*, is, I think, due to some subtleties in the Newtonian concept of uniform motion. According to the ordinary interpretation, uniform motion is a *state*. At first, this may seem self-contradictory. Is not motion *change* of place? How can a change be regarded a state? The mystery disappears if one keeps the concepts of velocity (which refers to a tendency) and change of place distinct.

That there has to be a distinction between velocity and change of place, can be seen from the fact that a velocity can either exist at a (mathematically) momentary point of time or endure for a time period, whereas a change of place *necessarily* takes some time. There can be a velocity but not a change of place in a momentary instant. Velocity and change of place, although in some way existentially dependent upon one another, are not identical aspects of things. This means that a thing which undergoes a *change* of place, may, during the same time, be in a *state* of uniform *velocity*.

Inertial motion involves two moments at one and the same time: change of place and velocity. It is a complex state of affairs simultaneously constituted by a state and a change. Actually, even in this it resembles Aristotelian self-change. As I remarked in relation to Aristotle, when an

actual change has a final cause, then this cause (tendency) exists as an actual state of the changing thing. In a self-changing entity there is a non-changing property (tendency) which brings forth the changes of the entity.[4]

If the distinctions between natural and artificial change and between efficient and final causality are applied to inertial motion, then inertial motion seems to be a natural change with a final cause. Obviously, an inertial motion has no external and no efficient cause. Does that, however, really imply that inertial motion can be regarded as having directedness and being caused by a final cause?

Let us see what a collision between two material heavy bodies looks like, first from an Aristotelian point of view and then from the Newtonian perspective. According to Aristotle, heavy bodies not affected by external forces move towards the center of the universe. When they are externally affected they none the less retain, independently of their actual movement, their *tendency* to move towards this center. Aristotelian heavy bodies have a specific directedness independently of all collisions and their actual direction of movement. They tend towards the mid-point of the universe whatever happens to them. This tendency belongs to the nature of heavy bodies, and it is their final cause. Their 'goal' is to be in rest in the center of the universe.

A Newtonian body in inertial motion moves of itself towards the points along a straight line, and the body has in this sense a specific directedness. If the body is pushed it will change direction, i.e. it will get a new specific directedness. The important thing in the present connexion, however, is the fact that the old direction is not retained even as a tendency. According to Aristotle, the nature of a heavy body gives it one specific directedness. According to Newton, the nature of a body with mass is such that it in and of itself maintains a motion in a certain direction, but the situation, not the nature of the body, determines this direction.

If, for a moment, we fancy that stones can have intentions, then the difference between the Aristotelian and the Newtonian point of view looks as follows. An 'Aristotelian stone' has one specific all-embracing intention, the project of its life, namely to come to rest in the center of the universe. A 'Newtonian stone' has no such project. Its intentions are changing and situation-bound.

Aristotelian self-change gets a *single* direction from the nature of the thing involved, whereas Newtonian self-change has, so to say, *multiple direction*. This, I think, is the real difference between Aristotelian and

Newtonian *causa sui*. I want to stress that inertial motion really is a kind of *causa sui*, even though it does not deserve the name 'final causality' in the sense 'one-goal causality'.

Most modern discussions of tendencies focus attention on dispositions and efficient causality. I have, however, tried to show that the concept of tendency is tied to that of *causa sui*. Furthermore, I claim that my 'argument from physics' makes it very probable that even a modern material ontology needs the category of *causa sui*. Aristotle's species of it, final causality, is perhaps not needed in physics, but the genus *causa sui* seems to be necessary. This, in turn, means that there is a kind of non-mental directedness which has to be taken into account in a definition of intentionality.

5. INTENTIONALITY

Now, with the former section as background, the problem of the difference between tendency and intentionality can be made more concrete. We can ask what the difference is between the directedness of a body in inertial motion and a person who walks along a straight line with an intention to walk in a straight line. Does the difference merely consist in the fact that the person, but not the material body, is conscious of the directedness? Does the world contain only one kind of directedness which has two modalities: physical and mental? Or, are there two radically different categories, tendency and intentionality, which, because of some superficial similarities, both deserve the epithet 'directed'?

Let us look at some possible definitions of intentionality,[5] and compare them with physical directedness. In the definitions which follow, the term 'entity' can denote anything whatsoever, concrete or abstract, simple or complex. Here is the first proposal.

D1 An entity has intentionality =def. the entity is directed towards another entity.

This is the definition I hinted at already at the beginning of this paper. As I have said, it is too wide. It applies to tendencies as well. Therefore, let us try a more Brentanist version:

D2 An entity has intentionality =def. the entity is directed towards an inexistent entity.

This definition has another flaw. If we opt for D2 we will make a

relational theory of (some) intentional acts false by definition. Since, like Kevin Mulligan and Barry Smith, I not only think such a theory is non-contradictory, but happen to believe in it, I cannot subscribe to D2.[6] However, a slight modification solves this particular problem. Let us look at D3.

D3 An entity is capable of intentionality =def. the entity is capable of being directed towards an inexistent entity.

According to the last definition, inertial motion cannot be regarded as being intentional, since such a motion has to be directed towards *existent* points in space. Inertial motion is not capable of being directed along an inexistent line. This fact, however, does not solve our *general* problem. Change of place is a specific kind of physical change. Do we have any reasons to believe that no other kinds of inertial changes can appear in physics? Let us see whether we at least can conceive of such a kind.

What, for instance, would *inertial change of electric charge* be like? An entity with such a kind of inertia would, when not affected by forces influencing the charge, have a constant change-of-charge-velocity. The entity would and in and of itself bring forward the corresponding actual changes of electric charge. Its change-of-charge-tendency would be directed towards *inexistent* electrical charges, i.e. towards some quantities of electric charge which are to come into existence in the near future. I think we can find no philosophical reasons to exclude such tendencies from physical theorising. This means that definition D3 (D2, too, by the way) is not acceptable. Like D1, it is too wide. It applies to some conceivable tendencies, not only to intentionality.

Let us try another line of thought. If an entity which has a certain tendency is not affected by counteracting forces, then the tendency *has to* realize itself. This means that, even though a tendency can be directed toward something actually inexistent, *a tendency can never point at something which is necessarily inexistent.* A tendency has to be directed towards something which is physically possible. This constraint, obviously, is not a constraint for intentional acts. Self-contradictory thoughts are possible, but self-contradictory tendencies are not. This gives the clue to our next attempt at a definition:

D4 An entity is capable of having intentionality =def. the entity is capable of being directed towards logically impossible entities.

Here we have a definition which does not encompass tendencies, but now other flaws appear. The reason why the definition D4 contains the phrase '*is capable* of having intentionality', instead of just 'has intentionality', is the fact that most intentional acts are not directed towards logically impossible entities. The phrase '*is capable* of having intentionality', unfortunately, gives rise to problems of its own. Assume that higher animals, or small children, can have perceptual (intentional) acts without being capable of thinking. Then, they are capable of intentionality but not capable of being directed towards self-contradictory entities. This means that D4 is extensionally too narrow. It says something essential about intentionality, but it does not characterize the category of intentionality.

6. TEMPORALLY EXTENDED ENTITIES

There is, however, a simple solution to our problem. A solution which has nothing to do with the kinds of entities which intentional acts and states may be directed towards. Instead, it brings in the notion of temporal extension. All four definitions above disregard temporal features. Tacitly, I would say, they presuppose that the directedness spoken of may be temporally punctual. Both when we think of tendencies and of intentional acts, we easily think of them as momentary. Let us now see what they look like during a temporal interval.

Once again we take a state of inertial motion as exemplifying the concept of tendency. In a momentary instant there is only a velocity with its directedness. There can be no change of place in such an instant, since a change necessarily takes some time. *In a time interval*, however, the velocity and the change of place of an inertial motion are existentially dependent upon each other.[7] If there is no change of place there can be no velocity, and vice versa. In other words: the directedness of an inertial velocity is, in a temporal interval, existentially dependent upon a change of place. This holds true for inertial changes in general, not only for change of place. A tendency, not counteracted by another tendency, is always, in a temporal interval, existentially dependent upon a corresponding change.

If, on the other hand, we look at an intentional act like a perception or a thought of something, then nothing similar appears. The directedness of such intentional acts is not existentially dependent upon changes, which, in turn, depend upon the intentional act. An intentional act may exist in

a momentary instant as well as endure for a time without any corresponding changes occurring. Of course, the intentional act itself may change, but that is irrelevant since it corresponds to the case where one tendency is exchanged for another. It does not correspond to the case where one particular tendency brings forth changes. I would like to propose a fifth definition:

D5 An entity has intentionality =def. the entity is (i) directed towards another entity, and (ii), there is in a temporal interval, no mutual existential dependence between the directedness and changes of the entity.[8]

Alas, even this proposed definition seems to have a counter-example, namely intentions. If a person really has a specific intention, he will necessarily act (i.e. produce changes) in order to realize this intention. We seem to have, contrary to D5, an intentional act which is existentially dependent upon connected changes. Let us, however, take a closer look at intentions.

7. THE DUALITY OF INTENTIONS

To have an intention is not only to have a mental representation, there has to be a tendency to act, too. An intention has to contain both intentionality and a tendency. This fact, however, does not imply that intentionality and tendency are identical. The right conclusion is that they are different *moments* in Husserl's sense, and that, consequently, intentions are complex states of affairs. In other words, an intention is constituted by *two* categorially different kinds of directedness: intentionality and tendency. There are some well known observations which support this 'dual aspect theory of intentions'. One may (a) be mistaken about one's own intentions, and (b) one's intentions are not always put into action.

(a) If an intention is made up of both an intentional state and a tendency, two possibilities arise. Either these moments have the same directedness or they point in different directions. The latter fact obtains when one is mistaken about one's real intention. One is then not mistaken about one's intentional state, but mistaken about one's presumed corresponding tendency to act.

(b) In other cases, although one is not mistaken about one's intention, one none the less never tries to realize it. These cases can easily be

accounted for in terms of tendencies. A tendency can have counteracting tendencies, which means that a tendency can exist without actually bringing forth the corresponding changes. Buridan's ass does not move, but it has actual tendencies to move. Such facts are impossible to account for if intentions are identified with intentional acts or states.

An intention is constituted by a tendency and an intentional state even when the intention (or, more correctly, its moment of intentionality) is directed towards something physically or logically impossible. A man trying to create a *perpetuum mobile* has a tendency to act. And often he acts. He tries to build a physically impossible machine. Similarly, a man who believes in the existence of square circles may try to draw such figures again and again. In cases like these, the intentional state of the intention is directed towards something impossible, but the tendency of the intention is directed towards something possible. If this were not the case, a man with an impossible intention could not, contrary to our experience, do anything in order to realize his impossible vision.

I think these remarks suffice to show that intentions cannot be used as counter-examples to D5. This being so, I claim that the definition D5 captures the essence of the category of intentionality.

8. FORMAL ONTOLOGY TODAY

Now, at last, I come to formal ontology. In order to make clear the difference between tendency and intentionality, I needed two concepts from formal ontology, that of 'moment' and that of 'existential dependence'. The latter is used in the final definition of intentionality, and without the former I could not have stated my thesis about the duality of intentions. Most interesting, however, from the point of view of formal ontology, is the fact that I also needed the concept of temporal extension. Traditional accounts of formal ontology leave this notion out. All part-whole relations discussed, as well as the concept of existential dependence itself, seem to be regarded as having no essential relation to time. In this respect, of course, formal ontology is similar to formal logic. The question is: should it be?

Some of my claims about inertial motion can be summed up by saying that an inertial motion is a complex state of affairs which makes the following two statements true:

(a) In a temporal interval, inertial motion contains two mutually

existentially dependent moments, velocity and change of place.

(b) In a momentary instant, inertial motion is identical with its velocity.

Together, statements (a) and (b) imply that velocity is part of inertial motion in another sense than that in which the change of place is part of inertial motion, and that to make this sense clear one needs the distinction between temporal interval and momentary instant. Or, one needs at least a distinction which corresponds to that between a line (temporal extension) and a point (momentary instant). Points seem not to be part of a line in the same way as the extended line-parts are parts of it. This difference in part-whole relationships is, I think, something for formal ontology to explore.

9. SUMMARY

1. Even modern physics uses implicitly an Aristotelian category of *causa sui*, although not the (one-goal) final causality of Aristotle.

2. Ontological systems ought to incorporate the category of tendency, as well as that of intentionality.

3. The existence of tendencies, with their kind of directedness, necessitates a more precise characterization of intentionality.

4. The true contrast between tendency and intentionality makes some peculiar part-whole relationships visible. Relationships which have not, so far, been given adequate attention within formal ontology.

NOTES

[1] This revival of the concept of tendency starts with Anscombe & Geach, *Three Philosophers* (Blackwell: Oxford 1961) and continues with R. Harré's *The Principles of Scientific Thinking* (Macmillan: London 1970) and (together with E.H. Madden) *Causal Powers* (Blackwell: Oxford 1975). It is further develolped by R. Bhaskar in A *Realist Theory of Science* (Leeds Books: Leeds 1975). The concept of propensity is primarily connected with Popper's so-called propensity interpretation of quantum mechanics; see e.g. "Quantum Mechanics Without 'The Observer'", in Bunge (ed.) *Quantum Theory and Reality* (Springer: Berlin 1967). Since my paper was originally written there has also appeared N.

Cartwright's *Nature's Capacities and Their Measurement* (Clarendon Press: Oxford 1989), as well as my own *Ontological Investigations* (Routledge: London 1989)

[2] See D.M. Armstrong & N. Malcolm *Consciousness & Causality* (Blackwell: Oxford 1984) pp. 149-53.

[3] As far as I know, Mario Bunge is the first philosopher who has noticed this; see his *Causality* (Harvard UP: Cambridge, Mass. 1959) pp.108-11.

[4] This implies that there is also a distinction to be made between *causa sui* and spontaneity. See my *Ontological Investigations* chapter 7.

[5] Those who refuse to speak of real definitions altogether, may take the proposed definitions which follow as merely some undefined kind of characterizations.

[6] See my *Ontological Investigations* section 13.5, especially note 14; Mulligan and Smith, 'A Relational Theory of the Act', *Topoi* 5 (1986), pp. 115-30; Smith, 'Acta cum fundamentis in re', *Dialectica* 38 (1984), pp. 157-78.

[7] For a fuller treatment of the notion of 'existential dependence' and part of its Husserlian background , see my *Ontological Investigations* chapter 9.

[8] There are some subtleties involved which cannot be discussed here (cf.note 7). I have to write *mutual* dependence, since an intentional state is, in all probability, *one-sidedly* dependent upon changes in the brain.

[9] I use the concept of formal ontology as it is used in B. Smith (ed.) *Parts and Moments. Studies in Logic and Formal Ontology* (Philosophia: Munchen 1982).

WOLFGANG LENZEN

LEIBNIZ ON PROPERTIES AND INDIVIDUALS

The Christian idea that GOD once created the entire world literally out of nothing does not sound very reasonable or intelligible to a modern analytic philosopher. Yet a 17th century forerunner, if not founder, of analytic philosophy, Gottfried Wilhelm Leibniz, apparently did believe in this doctrine or, somewhat more exactly, in the minimally weakened doctrine that GOD created the world out of nothing plus one. Ten years ago the Stadtsparkasse Hannover brought out the following commemorative coin:

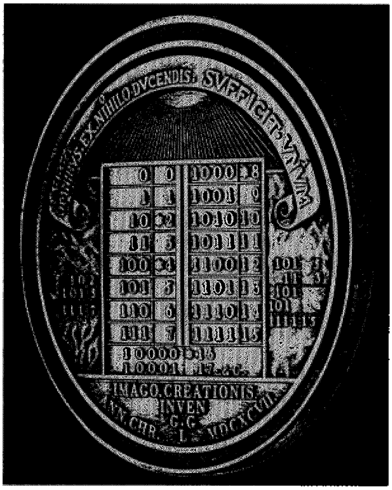


Figure 1.

In the middle of the coin we see a table with the beginning of the binary number system, enframed by samples of elementary arithmetical operations. On the top of the coin one can read the Leibnizian dictum "Omnibus ex nihilo ducendis sufficit unum", which may be translated as 'In order to produce everything out of nothing one [thing] is sufficient'. And the whole picture is said to present a "imago creationis" is, i.e. a picture of the creation.

Now, a sceptic will reply that the Hanover savings bank grossly misunderstood Leibniz's intentions. After all, the diagram only illustrates the by now well-known fact that the set of natural numbers can be built up from just two elements, namely from the numerals 0 and 1. Moreover, since Leibniz used to refer to the number 0 by the latin word 'nihil', the dictum 'Omnibus ex nihilo ducendis sufficit unum' can alternatively be translated as saying 'In order to produce every [number] from 0, [the number] 1 is sufficient'. Thus one might suspect that the bank, which is used to reducing everything to money, i.e., to numbers, quite mistakenly charged Leibniz with holding the Christian view of the creation of the *world* while in fact the famous 17th century mathematician stuck only to the modest claim that the *world of numbers* can be created from nothing plus one.

In 1697, however, Leibniz himself had painted and published the following picture:

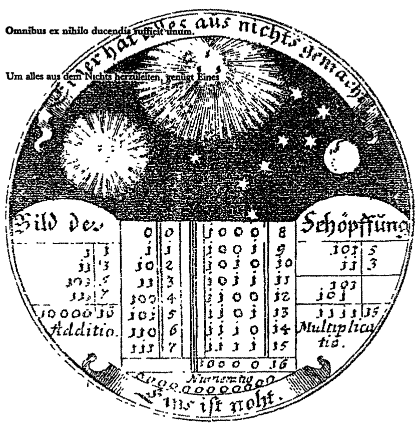


Figure 2.

Again we are told to see a "Bild der Schopffung", a picture of the creation. On top one can read "Einer hat alles aus nichts gemacht": One [namely GOD] has made everything out of nothing; and at the bottom the ambiguous statement "Eins ist noht" is added which can mean either that one *thing* or that the *number* one is necessary. In what follows I want to argue that Leibniz did not just have the trivial arithmetical interpretation in mind, but rather the full Christian doctrine of the creation of the world. Somewhat more exactly: Leibniz thought it possible for GOD to construct the *idea of the world* exclusively out of the idea of nothing and the idea of one in 7 steps:

The Logical Creation of the World

1) Starting with the numerals 0 and 1, one first obtains the set of *natural numbers*;
2) Each of these numbers is interpreted as representing (or being characteristic of) a specific *primary concept*;
3) By way of logical combination (conjunction and negation) the larger set of *general concepts* is obtained;
4) *Individual-concepts* (the "ideas" corresponding to individuals) will then be defined as maximally consistent concepts;
5) Among the set of all possible individuals the relation of *compossibility* is introduced;
6) *Possible worlds* are defined as certain maximal collections of pairwise compossible individuals;
7) *The real world* is distinguished from its rivals by being the most numerous (and, perhaps, also in some other respect the best) of all possible worlds.

Figure 3.

This 7-fold idea seems to underlie Leibniz's often-quoted remark: "Cum Deus calculat, fit mundus" which I would like to translate somewhat freely as saying that 'By way of GOD's calculations the world is created'.

In a recent book on "The System of Leibniz's Logic"[1] I have tried to give a formal reconstruction of this ambitious ontological enterprise. For reasons of space, however, I shall here concentrate on steps 3 and 4, which deal with Leibniz's conception of properties and of individuals. If we assume as given a denumerable set of concept letters A,B,C, ..., the main

elements of the algebra of concepts can be displayed in the following diagram:

The Algebra of Concepts

Operator	Symbol Leibniz	Set-theoretical Interpretation
Identity	$A = B$ $A \infty B$: eadem	Identity
Conjunction	AB $A + B$	Intersection $A \cap B$
Negation	$\overline{A}$ Non A	Complement
Containing	$A \subseteq B := (A = AB)$ A est/continet B	Inclusion $A \subseteq A$
Being Contained	AiB := BcA A inest ipsi B	Comprising $B \subseteq A$
Possibility	$P(A) := A \not\subset \overline{A}$ A est Ens/res	Non-emptiness $A \neq \Phi$
Empty Concept	$O := \overline{A\overline{A}}$ Nihil	Universe of discourse $\underline{U}$
Disjunction	$AvB := \overline{\overline{A}\,\overline{B}}$ commune	Union $A \cup B$
Communication	$C(A,B) := AvB \neq 0$ communicantia	Non-exhaustive-ness: $A \cup B \neq \underline{U}$
Subtraction	$A\text{-}B := Av\overline{B}$ $A \ominus B$	$A \cup \overline{B}$

Figure 4

The relation of *identity* or coincidence of concepts hardly needs any comment. Leibniz either symbolized it by the usual sign of equality or by the modern symbol of infinity. Often, however, Leibniz also referred to the identity of concepts informally by simply calling them the same ("idem", "eadem").

The operator of *conjunction* is here symbolized by mere juxtaposition. This follows Leibniz's main usage. Only in the fragments pertaining to the

so-called Plus-Minus-Calculus did he favor the mathematical '+'-sign to express a conceptual conjunction. The set-theoretical interpretation is straightforward: the extension of the conjunctive concept A-and-B is the set of all individuals that fall under both concepts, i.e., which belong to the intersection of the extensions of A and B.

The third primitive element of the algebra of concepts – and, by the way, the one with which Leibniz had notorious difficulties – is *negation*. Here it will be symbolized by drawing a line above the concept letter. Leibniz himself expressed the negation of a concept by means of the same word he also used to express propositional negation, namely the word 'non'. The extensional interpretation of Non-A, again, is rather straightforward. An individual x falls under the negative concept Non-A iff it fails to fall under the concept A; hence the extension of Non-A is just the set-theoretical complement of the extension of A.

Let us now consider the definable operators! The relation of conceptual *containment* is fundamental. My symbol 'c' derives from the Latin word 'continet' which Leibniz used on a par with the word 'est'. Here it is very important to observe that the set-theoretical interpretation of the intensional operator of containing always refers to the extensions of the concepts. Thus the proposition that concept A contains concept B has to be interpreted as saying that the extension of A – i.e. the set of all individuals that fall under concept A – *is included* in the extension of concept B. The underlying law of reciprocity was explained by Leibniz in the following passage from the *New Essays Concerning Human Understanding*:

> The common manner of statement concerns individuals, whereas Aristotle's refers rather to ideas or to universals. For when I say 'Every man is an animal' I mean that all the men are included amongst all the animals; but at the same time I mean that the idea of animal is included in the idea of man. 'Animal' comprises more individuals than 'man' does, but 'man' comprises more ideas or more attributes ... one has the greater extension, the other the greater intension.[2]

Accordingly, the converse relation that concept A *is contained in* concept B – AiB, or, as Leibniz puts it: "A inest ipsi B" – amounts to the set-theoretical condition that the extension of A comprises the extension of B or that the latter in contained in the former.

Closely related to the operator of negation is that of *possibility* or self-consistency of concepts. Leibniz expressed it in various ways, e.g. by saying 'A est possible' or 'A est ens' or 'A est res'. Sometimes the self-

consistency of A is expressed very elliptically by just saying 'A est', i.e., 'A is'. I use the capital P to abbreviate 'possibility'. Now, the extensional interpretation stated in the diagram might appear too weak. It is required there that concept A is possible if and only if it has a non-empty extension. But evidently there are certain concepts such as that of a unicorn which seem to be empty but which may nevertheless be regarded as possible, i.e. as not involving a contradiction.

It has to be observed, however, that the domain of individuals underlying the extensional interpretation of concept logic according to Leibniz is not taken to consist of actually existing objects only, but rather to comprise merely possible individuals as well. Given this specific understanding of the semantical requirements as referring to the set of all possible individuals, the non-emptiness of the extension of A then amounts to the condition that there is at least one possible individual that falls under the concept A. This, however, is both sufficient and necessary for A to be self-consistent. For if a certain concept A is possible then there is in fact among the vast collection of all possible individuals in particular at least one such object x which has the property A.

The next element of the algebra is the *empty concept* 0. Leibniz referred to it as 'Nihil', sometimes abbreviated by 'N'. The concept of Nihil or Nothing may easily be shown to be the same as the concept of a tautology and hence to have a universal extension.

Next we have a conceptual operator that would nowadays be called 'disjunction', but which appears in Leibniz's writings only in disguise, namely as the so-called "*commune*" of A and B, i.e., the common part of the concepts A and B or as what "A and B have in common". Leibniz never realized that the "commune" of A and B is the same as the disjunction, i.e. the 'or'-connection of both concepts. Nor did he therefore recognize that the "commune" of A and B can be defined as the negation of the conjunction of the individual negations of A and B. Anyway, this relation does obtain and it entails the condition of extensional interpretation as formulated in the diagram.

The next operator is the relation of *communication* of two concepts. As Leibniz puts it "A et B sunt communicantia" or "A et B habent aliquid commune" (A and B have something in common). This relation holds iff there is some concept C (different from 0) such that C is contained both in A and in B. This in turn is the case iff the commune or disjunction of A and B does not equal the empty concept, 0. From the latter, defining equivalence one obtains the semantical requirement stated in the diagram.

Finally, we have the *Minus-operation*, i.e. conceptual subtraction A-B. Leibniz's introduction of this operator was a kind of experiment. He simply tried to state the laws of conceptual or, as he called it, of "real" subtraction in direct analogy to the laws of arithmetical subtraction. Eventually he found a set of theorems which allows one to define A-B as the commune of A and Not-B, although Leibniz himself failed to realize this. Anyway, this definability justifies the semantical requirement listed in the diagram.

Let us now have a brief look at the Axioms of L1.

Axioms of L1

Ax1	AcA "B is B" - GI^3, § 37
Ax2	(AcB & BcC) $\rightarrow$ AcC 'If A is B and B is C, A will be C" o.c. § 19
Ax3	(AcB & AcC) $\leftrightarrow$ AcBC "That A contains B and A contains C is the same as that A contains BC", o.c., § 35
Ax4	"Not-not-A = A" o.c., § 96
Ax5	$A \neq \overline{A}$ "A proposition false in itself is 'A coincides with not-A'" o.c., § 11
Ax6	AcB $\leftrightarrow$ $\overline{B}c\overline{A}$ "'A is B' is the same as 'Not-B is not-A'" o.c., § 95
Ax7	AcB $\leftrightarrow$ $\neg$ P(A$\overline{B}$) "If I say 'A not-B is not [possible] this is the same as if I were to say ...'A contains B'" o.c., § 200.

Figure 5

Ax1 and Ax2 show that the relation of containment is reflexive and transitive. Ax3 is the only axiom for conjunction, and it establishes a connection between conceptual conjunction on the one hand and propositional conjunction on the other. The negation-operator is axiomatized by means of 3 principles: the law of double negation Ax4; the law of consistency Ax5; and the well known principle of contraposition, Ax6. Finally, the important Ax7 says that concept A contains concept B

iff the conjunctive concept A Non-B is impossible. This principle also characterizes negation, though only indirectly, and it completes the calculus L1. Let me mention in passing that Leibniz's algebra of concepts is provably equivalent or isomorphic to Boole's algebra of sets and that in this sense Leibniz invented the algebra of sets about 160 years before Boole.[4]

What do we learn from Leibniz's algebra of concepts for the issue of our conference? First of all, I do not see any reason to deny the existence of some kinds of properties. If the basic concepts are thought of as expressing properties, why should we deny the same status either to the negation of such a property or to the conjunction of two such properties. In this sense, then, there exist conjunctive and negative properties as well as disjunctive properties, empty properties, subtractive properties and what not. Contrary to what Leibniz thought, however, there is no such thing as a *privative* property. In the fragment C 264-270[5] Leibniz had attempted to classify all concepts into three subclasses of privative, semi-privative and positive concepts in the following way:

> If you subtract from some C a B which is not in it, then the rest A = C-B will be semi-privative; and if you adjoin to it some D, then E = D + A means that D and C have to be put in E, but from D B has to be subtracted first ... [If one has] E = L-M and L and M have nothing in common; and if (the uncommunicating) L and M are both something positive, then E is *semi-privative*. If, however, M = 0, then E = L is *positive* ...; and if, lastly, L = 0, then E = -M, i.e. E is privative. (C 267/8).

A privative concept, then, is supposed to be "less than nothing",[6] i.e. Leibniz thought that when e.g. – M (= 0-M) is added (i.e. conjunctively conjoined) to the corresponding positive concept M, the result would just be 0! It is not very difficult to show,[7] however, that the assumption of such privative concepts satisfying Leibniz's principle

$$(0\text{-}M)M = 0$$

would be incompatible with the basic theorems of *L1*. As a matter of fact it may instead be proven that any concept of the form – M or (0-M) coincides with 0 itself.

In order to investigate the relation between properties and *individuals* (or between concepts in general and individual concepts in particular), we now have to take a look at the quantifier logic L2:

Fundamentals of L2

Q1 $\alpha(A) \vdash \exists y\alpha(y)$

"For any definite letter there can be substituted an indefinite letter ... i.e. one can put A = Y" GI, § 23

Q1a "If A = AB, there can be assumed a Y such that A = YB" C, 235, #10.

Q1b "If AB is C, it follows that AY is C: or, it follows that some A is C. For it can be assumed that B = Y." GI § 49

Q2 $\neg \exists Y\alpha(Y) \leftrightarrow \forall Y^* \neg \alpha(Y^*)$

Q2A $\neg \exists Y(AYcB) \leftrightarrow \forall Y^*(\neg AY^*cB)$

"It must be seen whether, when it is said that AY is B ... Y is not taken in some other sense than when it is denied that any A is B ... so that when it is said that no A is B, the sense is that it is denied that AY* is B" GI, § 112

Q3 $\bar{\forall} Y\alpha \rightarrow \exists Y\alpha$

" ... for Y* is Y, i.e. any Y will contain this Y" ibid.

Q4 $AcB \leftrightarrow \forall Y(YcA \rightarrow YcB)$

"'A is B' is the same as to say 'If Y is A, it follows that Y is B'" C, 260 #15

Q5 $A\not{c}B \leftrightarrow \exists Y(P(AY) \& AYc\overline{B})$

"To say 'A is not B' is the same as to say 'There is a Q such that QA is not-B ... [and] QA is possible" C, 261, #18

Q6 $I(B) \leftrightarrow P(B) \& \forall Y(P(BY) \rightarrow BcY)$.

"So if BY is [possible], and the arbitrary indefinite term Y is superfluous, then B is an individual" GI, § 72

Figure 6

The first quotation from § 23 of the GI contains what would nowadays be called the rule of introduction of the existential quantifier: if a proposition α is true for a certain concept A, then there is some concept Y for which α holds. The subsequent passages present two applications of this fundamental principle. The first is particularly interesting since it explicitly contains the quantifier expression "there can be assumed a Y such that" which was otherwise mostly suppressed by Leibniz.

Next we have Leibniz's anticipation in § 112 GI of the logical relation between the universal and the existential quantifier: if one denies that for

some concept Y 'AY is B', this means that for arbitrary Y* the proposition 'AY* is B' is false; and the following passage expresses the law Q3 according to which a universally quantified proposition entails its existentially quantified counterpart.

Theorems Q4 and Q5 are typical examples of Leibniz's use of indefinite concepts as universal and as existential quantifiers. The former says that concept A contains concept B iff every concept X which contains A also contains B. Q5 conversely maintains that concept A fails to contain concept B iff there is a concept Y which is compatible with A such that the conjunction AY contains Non-B. This, by the way, is another of the few passages where Leibniz explicitly adds the quantificational phrase 'there is some Y such that'.

Q6 formalizes the "completeness" condition for individual concepts if and formulated by Leibniz in the concluding sentence of § 72 GI. There he called a concept Y "superfluous" (with respect to concept B) iff (for every C) $BY = C$ entails that $B = C$. This condition may be simplified by just requiring that Y is already contained in B. Somewhat more exactly: B is an individual-concept if and only if B is self-consistent and – unlike other concepts – B is complete in the precise sense of already containing any concept Y with which B is *compatible* (i.e. for which P(BY) holds). Incidentally, this condition can be simplified by the formula $I(B) \leftrightarrow \forall Y(BcY \leftrightarrow B\not{c}Y)$.

With the aid of this definition one may introduce a new sort of quantifier ranging over individuals – or more exactly: over individual-concepts – as follows:

$$\text{Q7} \qquad \begin{aligned} \bigvee X\alpha &:= \exists X(I(X) \,\&\, \alpha) \\ \bigwedge X\alpha &:= \forall X(I(X) \rightarrow \alpha) \end{aligned}$$

The truth-conditions for these formulae show that it is legitimate to speak of quantifying over objects here. E.g., the proposition that concept A is possible, i.e. that at least one (possible) individual falls under the concept A, can be expressed intensionally by saying that at least one individual-concept X contains the concept A and this may be rendered as $\bigvee X(X \subset A)$. Similarly, a singular predication of the type 'Socrates is mortal' simply takes the form ScM – the individual concept of Socrates contains the concept of mortality!

Now modern analytic philosophers believed that this was an error. Thus Bertrand Russell once pointed out: "Traditional logic regarded the two propositions 'Socrates is mortal' and 'All men are mortal' as being of

the same form; Peano and Frege showed that they are utterly different in form. The philosophical importance of logic may be illustrated by the fact that this confusion – which is still committed by most writers – obscured not only the whole study of the forms of judgment and inference, but also the relation of things to their qualities".[8]

What did Peano and Frege, or what did set-theory and predicate logic, really show? Set-theory has shown that a singular predication can be analyzed by means of the ϵ-relation as having the form $A \epsilon B$, while a universal predication is most naturally represented by means of the relation of inclusion among sets as having the form $A \subseteq B$. Similarly predicate logic has shown that universal predications can be analyzed by means of quantifiers as having the form $\forall X(Ax \rightarrow Bx)$ while a singular predication is represented (*apparently* without the help of any logical operator) in the form of a functional application: $B(A)$. However, these formalizations of the two types of propositions do not as such disprove the Leibnizian conception according to which both singular and universal propositions may also be analyzed as having one and the same logical form, AcB.

To be sure, the relation that an individual A bears to a quality B is ontologically different from the relation that obtains between two universals or two properties A and B. Nevertheless it is unobjectionable to paraphrase predications about both types in concept logic by means of one and the same formula: A contains B; the only difference being that in the former but not in the latter case A is an individual concept. The extensional interpretation of concept logic confirms the correctness of Leibniz's treatment of singular and universal propositions. It follows from the definition of individual concepts that if A is maximally consistent then the extension of A is minimal, i.e. it constitutes just a unit set $\{a\}$ containing one element of the underlying domain of possible individuals. According to the law of reciprocity, the individual concept A contains another concept B iff the extension of the former, i.e. the unit-set $\{a\}$ is contained in the extension of the latter, and this evidently is the case iff the individual a itself falls under the concept B, i.e. belongs to the extension of B. Hence in full accordance with the results of Peano and Frege the Leibnizian analysis of singular propositions and universal propositions as involving one and the same connective of conceptual containment turns out to be absolutely correct.

To conclude let me point out some "defects" in the orthodox form of predicate logic. First, traditional logic makes clear that the modern way of

expressing singular predications is rather misleading. When 'John is clever' is symbolized in the functional form C(j), it is thereby suggested that the copula 'is' is not itself a logical operator but only part of the predicate 'is clever'. This, however, seems dubious. For after all the operation of "applying" the function or the predicate of being clever to a particular individual (or to the corresponding proper name) constitutes a *logical operation*. In view of the fundamental character of this operation, it would seem appropriate to symbolize it explicitly by some new logical sign, e.g. '∋' (to remind us of the link with set-theoretical "predication") instead of hiding it behind the functional brackets.

Second, it would seem advisable to admit also predicational operators such as predicate negation, predicate conjunction, etc. To be sure, there is no logical fault in eliminating these operators in favor of corresponding propositional operators. There is, e.g., no reason to deny that an individual x has the conjunctive property 'F and G' iff x has both properties, F and G. Similarly, I do not wish to maintain (as is done by those logicians who think it necessary to distinguish "outer" from "inner" negation) that x's having the negative property Non-F, is not equivalent to x's failure to have the property F. Yet one would want to have at least the possibility of expressing these relations in some such form as

$$\wedge x((FG) \ni x \leftrightarrow F \ni x \wedge G \ni x) \text{ and } \wedge x\,(F \ni x \leftrightarrow \neg(F' \ni x)).$$

NOTES

[1] *Das System der Leinbnizschen Logik*, Berlin (de Gruyter), 1990.

[2] GP 5, 469 = C.J. Gerhardt (Hrg.) *Die philosophischen Schriften von G.W. Leibniz* (Berlin, 1875-90), Vol. 5.

[3] = *Generales Inquisitiones de Analysi Notionum et Veritatum* (hrg. u. ubers. v. F. Schupp), Hamburg (Meiner), 1982.

[4] Cf. my paper "Leibniz und die Boolesche Algebra", *Studia Leibnitiana* 16 (1984), 187-203.

[5] = *Opuscules et fragments inédits de Leibniz* (ed. L. Couturat), Paris 1903.

[6] Cf. GP 7, p. 233, fn. where Leibniz explicitly claims: "Hinc detractiones possunt facere nihilum ..., imo nimus nihilo", i.e. 'By the way of subtraction nothing ..., indeed even less than nothing can be obtained'.

[7] Cf. W. Lenzen: "Arithmetical vs. 'Real' Addition: A Case Study of the Relations Between Logic, Mathematics, and Metaphysics in Leibniz" in N. Rescher (ed.): *Leibnizian Inquiries*, Lanham: University Press of America, 1989, 149-57.

[8] Quoted according to F. Sommers: "Frege or Leibniz?", in M. Schirn (Hrg.) *Studien zu Frege*, Vol. III (Stuttgart, 1976), p. 13.

INDEX OF NAMES

INDEX OF SUBJECTS

PHILOSOPHICAL STUDIES SERIES

Founded by Wilfrid S. Sellars and Keith Lehrer

Editor:

KEITH LEHRER, *University of Arizona*

Board of Consulting Editors:

Jonathan Bennett, Allan Gibbard, Robert Stalnaker, and Robert G. Turnbull

1. JAY F. ROSENBERG, *Linguistic Representation*, 1974.
2. WILFRID SELLARS, *Essays in Philosophy and Its History*, 1974.
3. DICKINSON S. MILLER, *Philosophical Analysis and Human Welfare*. Selected Essays and Chapters from Six Decades. Edited with an Introduction by Lloyd D. Easton, 1975.
4. KEITH LEHRER (ed.), *Analysis and Metaphysics*. Essays in Honor of R. M. Chisholm. 1975.
5. CARL GINET, *Knowledge, Perception, and Memory*, 1975.
6. PETER H. HARE and EDWARD H. MADDEN, *Causing, Perceiving and Believing*. An Examination of the Philosophy of C. J. Ducasse, 1975.
7. HECTOR-NERI CASTAÑEDA, *Thinking and Doing*. The Philosophical Foundations of Institutions, 1975.
8. JOHN L. POLLOCK, *Subjunctive Reasoning*, 1976.
9. BRUCE AUNE, *Reason and Action*, 1977.
10. GEORGE SCHLESINGER, *Religion and Scientific Method*, 1977.
11. YIRMIAHU YOVEL (ed.), *Philosophy of History and Action*. Papers presented at the first Jerusalem Philosophical Encounter, December 1974, 1978.
12. JOSEPH C. PITT, *The Philosophy of Wilfrid Sellars: Queries and Extensions*, 1978.
13. ALVIN I. GOLDMAN and JAEGWON KIM, *Values and Morals*. Essays in Honor of William Frankena, Charles Stevenson, and Richard Brandt, 1978.
14. MICHAEL J. LOUX, *Substance and Attribute*. A Study in Ontology, 1978.
15. ERNEST SOSA (ed.), *The Philosophy of Nicholas Rescher: Discussion and Replies*, 1979.
16. JEFFRIE G. MURPHY, *Retribution, Justice, and Therapy*. Essays in the Philosophy of Law, 1979.
17. GEORGE S. PAPPAS, *Justification and Knowledge: New Studies in Epistemology*, 1979.
18. JAMES W. CORNMAN, *Skepticism, Justification, and Explanation*, 1980.
19. PETER VAN INWAGEN, *Time and Cause*. Essays presented to Richard Taylor, 1980.
20. DONALD NUTE, *Topics in Conditional Logic*, 1980.
21. RISTO HILPINEN (ed.), *Rationality in Science*, 1980.
22. GEORGES DICKER, *Perceptual Knowledge*, 1980.

23. JAY F. ROSENBERG, *One World and Our Knowledge of It,* 1980.
24. KEITH LEHRER and CARL WAGNER, *Rational Consensus in Science and Society,* 1981.
25. DAVID O'CONNOR, *The Metaphysics of G. E. Moore,* 1982.
26. JOHN D. HODSON, *The Ethics of Legal Coercion,* 1983.
27. ROBERT J. RICHMAN, *God, Free Will, and Morality,* 1983.
28. TERENCE PENELHUM, *God and Skepticism,* 1983.
29. JAMES BOGEN and JAMES E. McGUIRE (eds.), *How Things Are, Studies in Predication and the History of Philosophy of Science,* 1985.
30. CLEMENT DORE, *Theism,* 1984.
31. THOMAS L. CARSON, *The Status of Morality,* 1984.
32. MICHAEL J. WHITE, *Agency and Integrality,* 1985.
33. DONALD F. GUSTAFSON, *Intention and Agency,* 1986.
34. PAUL K. MOSER, *Empirical Justification.* 1985.
35. FRED FELDMAN, *Doing the Best We Can,* 1986.
36. G. W. FITCH, *Naming and Believing,* 1987.
37. TERRY PENNER, *The Ascent from Nominalism.* Some Existence Arguments in Plato's Middle Dialogues, 1987.
38. ROBERT G. MEYERS, *The Likelihood of Knowledge,* 1988.
39. DAVID F. AUSTIN, *Philosophical Analysis.* A Defense by Example, 1988.
40. STUART SILVERS, *Rerepresentation.* Essays in the Philosophy of Mental Rerepresentation, 1988.
41. MICHAEL P. LEVINE, *Hume and the Problem of Miracles. A Solution,* 1979.
42. MELVIN DALGARNO and ERIC MATTHEWS, *The Philosophy of Thomas Reid,* 1989.
43. KENNETH R. WESTPHAL, *Hegel's Epistemological Realism.* A Study of the Aim and Method of Hegel's *Phenomenology of Spirit,* 1989.
44. JOHN W. BENDER, *The Current State of the Coherence Theory.* Critical Essays on the Epistemic Theories of Keith Lehrer and Laurence Bonjour, with Replies, 1989.
45. ROGER D. GALLIE, *Thomas Reid and 'The Way of Ideas',* 1989.
46. J-C. SMITH (ed.), *Historical Foundations of Cognitive Science,* 1990.
47. JOHN HEIL (ed.), *Cause, Mind, and Reality.* Essays Honoring C. B. Martin, 1990.
48. MICHAEL D. ROTH and GLENN ROSS (eds.), *Doubting.* Contemporary Perspectives on Skepticism, 1990.
49. ROD BERTOLET, *What is Said.* A Theory of Indirect Speech Reports. 1990
50. BRUCE RUSSELL (ed.), *Freedom, Rights and Pornography.* A Collection of Papers by Fred R. Berger. 1990
51. KEVIN MULLIGAN (ed.), *Language, Truth and Ontology.* 1992
52. JESÚS EZQUERRO and JESÚS M. LARRAZABAL (eds.), *Cognition, Semantics and Philosophy.* Proceedings of the First International Colloquium on Cognitive Science, 1992
53. O.H. GREEN, *The Emotions.* 1992